DOCUMENT RESUME

ED 177 478 CS 005 001

AUTHOR Cardish, Edna Catherine
TITLE The Relationship between Reading Content and Reading Comprehension among Ninth Grade Pupils in an Urban Area--Population 114,973.
PUB DATE 3 Jul 71
NOTE 202p.; Ph.D. Thesis, The University of Sarasota

EDRS PRICE MF01/PC09 Plus Postage.
DESCRIPTORS Disadvantaged Youth; *Minimum Competency Testing; *Reading Ability; *Reading Comprehension; Reading Research; *Reading Tests; Secondary Education; Standardized Tests; Student Experience; *Test Bias; Test Construction; *Urban Youth

ABSTRACT

One hundred fifty-one urban ninth grade students participated in an investigation of the effect of differentiated content in two congruent tests of reading. A congruent test was constructed by substituting information relevant to junior high school students within the parameters of a nonrelevant, standardized state test of reading. By maintaining identical format, grammatical structure, flow of thought, patterning of questioning, and readability level, the two tests were shown to be of equivalent difficulty. Comparison of subscores of boys and girls, blacks and whites, as well as eight individual classes showed that there was a general tendency toward greater achievement on relevant reading content, with no one group significantly exceeding another statistically. Other results indicated that the content of the standardized state reading test is basically alien to a number of urban children and may penalize them unfairly, and that a number of apparently noncompetent readers in junior high school may actually possess minimal reading ability within the framework of their experiential background. (Tabular material is included, and sample test questions, word lists, raw scores, and a list of interest clarifying questions are appended.) (AEA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED177478

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE RELATIONSHIP BETWEEN READING CONTENT AND READING COMPREHENSION AMONG NINTH GRADE PUPILS IN AN URBAN AREA - POPULATION 114, 973

by

Edna Catherine Cardish

Dissertation Advisor

Professor Eugene A. Guarino

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Edna Catherine Cardish

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the University of Sarasota

July 3, 1971

CS005001

ACKNOWLEDGEMENTS

First and foremost the author wishes to express sincere appreciation to Laurence University for providing a truly unique and stimulating educational experience. It was a privilege and an inspiration to be associated with both the staff and the members of the first official student body.

One who undertakes an experiment within a public school system can do so only if he has the cooperation of many people. This author has been fortunate to be associated with a system which provided that cooperation. Special thanks are extended to Dr. James T. Hepinstall, Superintendent of Schools; Miss E. Olive Jennings, Coordinator of Federal Projects; Miss Doris E. Bedell, Director of Urban Education; Mr. Edward Handron, Principal of Philip Livingston Junior High School; and to the guidance counselors and teachers at Livingston who assisted in the administration of the experiment.

Gratitude is extended to Dr. Eugene Guarino who, despite a full schedule at the State University of New York at Albany, graciously agreed to direct this study.

In addition, the author wishes to acknowledge her indebtedness to Mr. John Cosgrove for his many constructive suggestions; to Mr. John Bartley for the preparation of the illustrations; and to Mrs. Olga Ahnert for her patience, pre-

cise typing, and outstanding editorial ability.

Gratitude is also extended to individuals whose names appear in the bibliography and to the publishers who generously gave permission to quote from materials in their publications.

Finally, sincere thanks are due to my father, my close relatives, and my personal friends for their sustained support without which this study would have been impossible.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	viii
LIST OF FIGURES	ix
Chapter	
1. INTRODUCTION	1
THE PROBLEM	3
Statement of the Problem	6
Sub-Problems	6
THEORETICAL FRAMEWORK	11
Definitions	11
Basic Assumptions	22
SIGNIFICANCE OF THE STUDY	26
Affecting Test Change	26
Affecting Attitude Changes	27
Initiating Similar Research	28
Bearing on Current Problems	29
2. RELATED LITERATURE	30
CURRENT CONCERNS	30
Standardized Testing	30
Testing and Curriculum	35
Norm-referenced versus Criterion-referenced Measurements	36
Testing and Financing	37

Chapter Page

PROFESSIONAL OPPOSITION 40

PROFESSIONAL SUPPORT 46

3. DESIGN AND PROCEDURES 65

HYPOTHESIS 65

TEST MATERIALS 66

New York State Minimum Competence Test in Reading 66

Reading Ability Test 68

Congruency of Tests 72

Content of Tests 85

SUBJECT POPULATION 89

TEST PROCEDURES 90

4. PRESENTATION AND ANALYSIS OF THE DATA 91

TEST OF THE NULL HYPOTHESIS 91

Mean Differential 91

Coefficients of Correlation and Alienation 91

OTHER SUPPORTIVE DATA 97

Four-Fold Distribution of Compared Scores . 97

Sex and Race Mean Score Comparisons 99

Comparison of Individual Classes 100

Reliability of the Tests 110

Item Analysis of the Tests 111

Chapter Page

5. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS . . . 124

SUMMARY 124

Problem 124

Procedure 125

Results 126

CONCLUSIONS 127

RECOMMENDATIONS 129

BIBLIOGRAPHY 133

APPENDICES 141

LIST OF TABLES

Table		Page
1.	Dale-Chall Readability Rating of Paragraphs for the New York State Minimum Competence Test in Reading and the Reading Ability Test	76
2.	Reading Difficulty Range Chart for the New York State Minimum Competence Test in Reading and the Reading Ability Test	78
3.	Items Eliciting Best and Poorest Responses on the Minimum Competence Test and the Reading Ability Test	95
4.	Four-Fold Distribution of Compared Scores - New York State Minimum Competence Test in Reading - Reading Ability Test	98
5.	Mean Score Comparisons of Sex and Race Factors in Relation to Relevant Reading Material	101
6.	Analysis of Item #26 in Terms of High and Low Scorers	116
7.	Best Discriminatory Items on the Minimum Competence Test in Terms of High and Low Scorers	120
8.	Best Discriminatory Items on the Reading Ability Test in Terms of High and Low Scorers	120

LIST OF FIGURES

Figure		Page
1.	Meaning-Space Concept Chart	52
2.	Basic Correlates of Reading Achievement and Deficiency	62
3.	A Comparison of the Reading Content of the New York State Minimum Competence Test in Reading and the Reading Ability Test as Each Relates to the Realistic Life-Space of Children in Urban Areas	71
4.	Comparison of Congruent and Non-Congruent Factors in the New York State Minimum Competence Test (M.C.T.) in Reading and the Reading Ability Test (R.A.T.)	73
5.	Scattergram	93
6.	Frequency Polygons - Class A	102
7.	Frequency Polygons - Class B	103
8.	Frequency Polygons - Class C	104
9.	Frequency Polygons - Class D	105
10.	Frequency Polygons - Class E	106
11.	Frequency Polygons - Class F	107
12.	Frequency Polygons - Class G	108
13.	Frequency Polygons - Class H	109
14.	Item Analysis Percentage of Correct Answers	112
15.	Discriminating Power of Items on the Minimum Competence Test in Reading	118
16.	Discriminating Power of Items on the Reading Ability Test	119

Chapter 1

INTRODUCTION

Since 1965 a growing discrepancy has been observed among public school pupils in disadvantaged areas of Albany, New York. Children appear to read better each year but annual reading test scores fail to reflect this apparent growth. Scores continue in a disappointing pattern despite the fact that classroom evidence attests to the fact that more children are reading better than ever.

Six years of federal support, coupled with three years of urban education funding together with increasing local expenditures for reading improvement per se have created an atmosphere in which reading receives primary emphasis and observed classroom behaviors indicate that children can and do read better than test results indicate.

At the elementary level, for example, third grade children manifest a strong degree of mastery over language elements. The majority of them are reading up to grade level in various basal readers. A staff of reading assistants has been on hand since 1965 to reduce and/or eliminate the reading problems of elementary children as those problems were forming. Site visits to classes demonstrate that the children, as a whole, not only can and do read, but read with an apparent degree of authentic enjoyment. Despite these encour-

aging signs, grade three classes tend to score below expectancy on the New York State Pupil Evaluation Test in Reading.

At the secondary level, as well, the same discrepancy arises. In high schools, for example, reading scores continue in a similar, locked pattern despite the fact that paperbacks wear out rapidly, relevant books disappear, the demand for more books constantly exceeds the supply, teachers vie with each other for extra copies of student magazines, and the majority of children manifest behaviors which indicate that they possess competence in reading.

What is happening in Albany is occurring in other urban areas throughout the country. Reading programs have never been stronger due to the influx of various sources of supplementary funding. Despite this fact, both national and state reading test pictures are equally disappointing in spite of the substantial input of effort.

Several questions arise in light of these discrepancies, for example: Are supplementary funds having little, if any, effect? Must funding be doubled, or even tripled, in order to produce significant reading gains? Are the children being measured properly? Is there any aspect of the measuring devices themselves which may be deflecting scores? Are currently used reading devices valid, particularly for children who live in urban areas?

Questions concerning the impact of current or increased funding with the intent of improving reading may be answered through national reading surveys and reading test programs now in effect.

The question of the validity of current reading measures is the immediate concern of this study. Why is there a clash between the empirical evidence found in classrooms and the statistical data produced through annual testing? To make this a workable question, the study had to focus on one test at one grade level for one group of children. If it can be demonstrated that a designated test of reading is an invalid and/or inaccurate measure, the discrepancy between classroom performance in reading and reading test performance can be partially explained. If, in addition, replication of this experiment is made in other places and with other tests, the body of evidence produced might indicate that the national reading picture is not as bleak nor discouraging as that currently in vogue.

THE PROBLEM

Since 1962 the New York State Minimum Competence Test in Reading has been administered to high school pupils in order to differentiate between children who possess minimum competence in reading and those who do not. This test was selected for this study as a means of determining if the nature of the content of a test of reading has an effect on the level of reading ability achieved by junior high school children.

The Minimum Competence Test was first administered to seventeen hundred twelfth grade pupils in a random sample on Long Island and in upstate New York. Some 97 per cent of

this sample scored twenty-six points or higher out of a possible forty. In the fall of 1965 the test was administered to 227,920 ninth grade public school pupils; 83 per cent of this group scored twenty-six or higher. Of the public and non-public school ninth grade children combined, 85 per cent scored twenty-six or more correct answers. A score of twenty-six became the cut-off point for differentiating between competent and non-competent readers. The State Test Manual[1] advises that the test results be used in two ways: namely, to certify a level of reading achievement necessary for graduation and to determine the extent of need for remedial reading programs which will allow children to reach a level of reading competency prior to the time they graduate.

The Minimum Competence Test was selected as the basis of this experiment for two main reasons. First, it is administered annually throughout New York State and results are used to judge general reading ability of children in various parts of the State. Cities are compared with one another and both State and local administrators examine results to evaluate the reading status of children in districts. A test with such strong implications merits close scrutiny. Secondly, in order to test a theory of the impact of content on reading ability, a new test had to be constructed to match the Minimum Competence Test. The structure of the Minimum Competence

---

[1]New York State Education Department, Manual of Directions for the New York State Minimum Competence Test in Reading (Albany: New York State Education Department, 1969), p. 5.

Test is less complex than that of many typical tests of reading ability; therefore, it was feasible to attempt to match this test with a new counterpart for experimental purposes.

The New York State Minimum Competence Test in Reading is mandated for ninth grade children and is administered early in October of each year. For many children, perhaps the majority, the test presents no substantial difficulty; for advanced placement groups it could hardly be considered an obstacle. For a number of children, however, the content of the test itself may be a barrier to their scoring minimum competence in reading. It is this low achievement group, the so-called disadvantaged, who are the concern of this study. These children may be penalized by the fact that on the Minimum Competence Test they are asked to read content which, for the most part, falls outside the realm of their experimental background.

The New York State Minimum Competence Test in Reading is not a fully standardized nor normative test. Is is currently being considered for revision and there is time to initiate changes in the test in accord with the findings of this study should those findings prove significant.

Although it was necessary to construct a new measuring tool in order to test a theory of the impact of content on reading ability, this study does not propose to develop a replacement for the Minimum Competence Test. The matching test, entitled the Reading Ability Test, was devised solely for the purpose of testing an hypothesis. It is not proposed

as an improved nor ideal substitute for the Minimum Competence Test, but merely as a vehicle for providing an insight into the relationship, if any, between content and comprehension.

Statement of the Problem

This study was designed (1) to test the content validity of the New York State Test of Minimum Competence in Reading as a suitable measure of reading ability among urban children, (2) to demonstrate that the content of the New York State Test of Minimum Competence in Reading is basically alien to a number of urban children, and, as such may penalize them unfairly, (3) to demonstrate that a number of apparently non-competent readers in junior high school may actually possess minimal reading ability within the framework of their experiential background, and (4) to establish data which may be instrumental in the revision of the New York State Test of Minimum Competence in Reading as well as in other tests of reading which share a similar limitation.

Sub-Problems

Test construction and administration. In order to determine if relevant content does indeed have a positive effect on comprehension, it was necessary to devise a test which tightly paralleled the New York State Test of Minimum Competence in Reading but presented the child with a new body of relevant reading material. Such a test had to share with its counterpart similar format, sentence flow, train of

thought, degree of readability, and pattern of questioning. It had to substitute foreign concepts with concepts more likely to fall within the experiential range of the children to be tested. Using Havighurst's[1] life-space concept, the new test had to incorporate not only an external, structural congruency, but a concept congruency as well so that test content could fall within a realistic life-space of the urban child. To arrive at this kind of reasonable content, it was necessary to use urban children themselves to identify concepts with which they were most familiar. Through a summer Communications Course conducted three years ago for another purpose, it was possible to tap this pupil source of information.

After designing the Reading Ability Test as one which was congruent with the New York State Test of Minimum Competence in Reading, it was necessary to match its readability level so that although differentiated by content, the two tests would share an identical degree of reading difficulty. An appropriate formula, the Dale-Chall[2] Formula for Predicting Readability, was selected and applied to each paragraph on both tests. Factors of sentence length, word count, and number of unfamiliar words were manipulated until all eleven

---

[1]Robert J. Havighurst and Bernice L. Neugarten, Society and Education (Boston: Allyn and Bacon, Inc., 1962), p. 206.

[2]Edgar Dale and Jeanne S. Chall, "A Formula for Predicting Readability," Educational Research Bulletin, I (1962), pp. 1-9, 11-28.

matching paragraphs on both tests proved of equal difficulty statistically.

In addition, the Reading Ability Test had to be placed through a number of revisions based on the findings of early item analysis results; poor matches and defective items had to be identified and corrected.

It was also necessary to check the notion of one test as being congruent with another. Since no evidence of any similar tool could be found, this step was, of necessity, a subjective one. A panel of experts consisting of nine professionals from the fields of reading, sociology, psychology, and secondary education were consulted as to the advisability of the new measure. Members were asked to rate the congruency of each paragraph and make recommendations for revisions where such were deemed necessary. These revisions had to be worked into the test proper without disturbing the readability levels previously determined.

Finally, suitable methods of dealing with the results had to be determined early in the study so that proper test conditions could be set up and appropriate data gathered.

On the completion of the design of the Reading Ability Test, it was necessary to have the two tests prepared in identical format together with two sets of appropriate and matching answer sheets. Permission was obtained from the State Education Department to have the Minimum Competence Test in Reading put into typewritten form so as to match the typed form of the Reading Ability Test.

Early in September of 1970 tests were scheduled for October administration at Livingston Junior High School in Albany, the target school for this experiment. Guidance counselors and teachers had to be briefed in advance as to the proper administration of the experiment.

With testing completed, scoring, item analysis, and both descriptive and inferential statistical procedures had to be conducted in order to provide the necessary basis for interpretation of the results.

Interdisciplinary investigation. Since this study considered the affective response of the child to a test situation in addition to his cognitive response, it was necessary to extend the investigation beyond the confines of education itself; findings of psychologists, sociologists, and anthropologists were, of necessity, a consideration of this study. If it is no longer sufficient for the educator to have as a sole objective of education the cognitive development of the child as pointed out by Crary[1] and Leacock,[2] it is, by extension, no longer sufficient to test a child without consideration of his affective nature. The child, as a whole person, is being tested; he brings to the printed word what the word means to him, what feelings, if any, he

---

[1]Ryland W. Crary, Humanizing the Curriculum (New York: Alfred A. Knoph, 1970), p. 16.

[2]Eleanor B. Leacock, Teaching and Learning in City Schools, A Comparative Study (New York: Basic Books, Inc., 1969), p. xiii.

associates with it, and the testing conditions under which he is asked to function in response to it. To confront a child from one environmental set of concepts with test items, values, and interests from another and ask him to cope successfully with that confrontation may be to test him unfairly. As Crary points out:

> The cultural or experience bias in both intelligence and academic tests has been apparent for some time--projects for reconstructing the experience of the so-called culturally deprived are really vindications on a major scale of the necessity for method.[1]

An assumption which seems to underlie many such tests of reading is that the mechanical linkage between words and objects developed through years of schooling makes any area of curriculum a sufficient base on which to construct test items. Vygotsky[2] cites research which destroys this assumption. He demonstrates that conceptual development is vitally related to processes rather than merely verbal associations. Furthermore, concept formation is a slow process which begins to emerge around the time of puberty. Prior to age twelve, thought concepts which are not based directly on perception put demands on children that exceed their mental capabilities. Such observations give rise to the need for investigation of how the testing process should realistically parallel the concept-formation-process.

---

[1]Crary, p. 24.

[2]Lev Semenovich Vygotsky, Thought and Language (Cambridge, Mass.: Massachusetts Institute of Technology, 1962), pp. 53-55.

THEORETICAL FRAMEWORK

Definitions

Congruent test. The term, congruent, has been adopted to describe the Reading Ability Test the administration of which formed the basis for this experiment. The descriptor, congruent, was utilized because of the lack of a testing term adequate to describe this kind of test. Although it is a type of parallel to the New York State Minimum Competence Test in Reading, the word parallel, as used to describe a companion test, has a different connotation indicating the rearrangement of the same content at the same level of difficulty. There is no evidence of a parody-type test as that constructed by this study which retains similar structure but substitutes new content within that structure.

The Reading Ability Test shares with the New York State Minimum Competence Test in Reading similar grammatical structure, the same flow of ideas, identical readability paragraph levels according to the Dale-Chall[1] formula of readability, and a similar pattern of questioning.

The notion of congruency was proposed to a panel of professionals in the field of reading, psychology, sociology, and secondary education. Panel members accepted the notion of congruency and rated paragraphs according to this criteria.

---

[1]Dale and Chall, pp. 1-9, 11-28.

Congruent content. The design of this experiment was not only to construct a test which would, in a sense, fit into the structure of another test, but to choose content for the new test which would coincide with the values, interests, and experience level of urban children. It was contended that a fair test of minimal reading ability should be written within a realistic range of knowledge and experience. Aaron points out that

> In order to get meaning, the student must be able to bring some meaning to the material so that he will be able to see the similarities between what he knows and what he reads . . . the processes which are involved in concept formation cannot be inferred from mental age or vocabulary, but rather, from the cumulative experiences of the individual that have had meaning for him.[1]

This study proposed that alien concepts interfere significantly with the reaction of socio-economically deprived children to the printed word.

Chapter 3 will demonstrate how the items on the New York State Minimum Competence Test in Reading tend to fall outside realistic experiential limits of a number of urban children. Aside from a chance acquaintance with typical items on the Minimum Competence Test, through social studies or science classes, most of the concepts presented on the State Test stand at a distance from the life-space of the urban child. The Reading Ability Test, on the other hand, draws content closer to that life-space allowing the child to

---

[1] Ira E. Aaron, "An Overview of Teaching Comprehension," Teaching Comprehension Skills (Atlanta, Ga.: State Education Department, 1964), p. 14.

bring more meaning to what he is asked to read in anticipation of increasing the degree of his achievement in reading.

Since people do not share identical developmental experiences, it would be impossible to design a test which falls completely within the precise range of a specific child's background; however, general interests and values of junior high school urban children are identifiable and a test writer may compose items within a realistic range of relevance. It is this type of substitutive content which is identified in this study as congruent content.

Urban children. The adjective, urban, as a descriptor of children could be objected to on the basis of the fact that children are the same the world over despite the accident of geographical birth. There are two reasons why children in this study have been referred to as urban. A sampling of children residing in a metropolitan area will include many so-called disadvantaged children; however, culturally disadvantaged is rejected for use in this study as a misleading term. Cohen[1] points out that a child from an ethnic group may be rich in his own culture but socially disadvantaged in terms of American culture. It is social disadvantagement which restricts his freedom to choose from among alternatives which control his destiny. Furthermore, a recent article in

---

[1]S. Alan Cohen, Teach Them All to Read: Theory, Methods, and Materials for Teaching the Disadvantaged (New York: Random House, 1969), p. 17.

the Saturday Review stated that "If there are any culturally disadvantaged ghettos, they are probably found in white suburbs . . . the ghetto may do more for the arts than the arts will do for the ghetto."[1] Therefore, the term disadvantaged is rejected as inappropriate for this or any study.

The second reason why urban children is preferred as a descriptor of children in this study is that although children are indeed alike, their language development is influenced in large measure by the concepts, values, and interests derived through their immediate environment. Geographic conditions determine cultural differences; these differences are characterized by a different view of life's problems, a different style of living which results in a different orientation to the printed page.[2] Therefore, children who live in large cities, attend large urban high schools, and reside within the general area referred to as the "city proper" are the object of this study and for purposes of identification will be referred to as urban children.

Foreign and familiar concepts. The crux of this study deals with the substitution on a test of reading of one set of concepts with another. It is necessary to differentiate

---

[1]Paul Woodring and others, "Education in America," Saturday Review, July 18, 1970, p. 41.

[2]Roger W. Shuy, "A Linguistic Background for Developing Beginning Reading Materials for Black Children," Teaching Black Children to Read (Washington: Center for Applied Linguistics, 1969), p. 118.

between foreign and familiar concepts; concepts are judged as to the degree of familiarity in relation to cultural background. Culture is defined by Lowry[1] as the acts, objects, ideas, and sentiments which constitute the basis of one's life. Therefore, familiar concepts, as far as this study is concerned, are those which actually exist within the surroundings and consequently within minds of urban children; content for the Reading Ability Test was suggested by, and grew out of, the real ideas, feelings, beliefs, expressions, and manifested interests of urban children themselves.

Foreign concepts, on the other hand, are basically ideas which urban children lack, or at least, fail to express. It may be objected that concepts found on the New York State Minimum Competence Test in Reading are not alien because they have been taught in social studies or science classes. This is likely the case; however, the teaching of such concepts is usually done verbally and, according to Piaget, "You cannot teach concepts verbally; you must use a method founded on activity."[2]

Familiar concepts, in this study, refer to more than verbal meaning; they refer to what Kwant[3] calls "lived meaning;"

---

[1]Ritchie P. Lowry and Robert P. Rankin, Sociology, The Science of Society (New York: Charles Scribner's Sons, 1969), p. 98.

[2]Elizabeth Hall, "A Conversation With Jean Piaget," Psychology Today, III (May, 1970), 30.

[3]Remy C. Kwant, Phenomenology of Language (Pittsburgh: Duquesne University Press, 1965), p. 39.

he speaks of the compenetration of language and existence. It is this compenetration which the Reading Ability Test seeks to establish; namely, the mingling of real life experiences with real-life-related reading activities.

Since an understanding of familiar concepts in terms of the urban child is crucial to an understanding of this study, one further illustration will be drawn. Although Warner[1] worked with children in a geographically alien culture, she discovered that the degree of familiarity of concepts formed the basis of her success in teaching Maori children to read. She developed a system of "organic reading" which proved that children can and do read well within the structure of their experiential background. One of her conclusions was that "The distance between the content of their (children's) minds, however, and the content of our reading books is nothing less than frightening."[2] This statement is amended by this study to decry a similar and equally shocking discrepancy between the content of an urban child's mind and the content of reading tests by which he is measured and labelled as a competent or non-competent reader.

Reading comprehension. Reading comprehension is an apparently simple concept implying the ability of an individual to extract meaning from printed symbols. However, when

---

[1]Sylvia Ashton Warner, The Teacher (New York: Simon and Schuster, 1963), p. 27.

[2]Warner, p. 69.

reading and language experts define comprehension, simplicity gives way to complexity and confusion may result. In fact, actual disagreement exists among experts as to the true meaning of comprehension;[1] definitions range from the simplest stated above to criterion-based definitions which identify a set of skills the application of which is said to constitute comprehension.[2] An extended definition which illustrates how difficult it is to capture the true meaning of the term is the following:

> Good reading is far more than recognizing words in print, far more than mere recalling main ideas and related details. Good reading demands active participation of the reader in reacting to the ideas set forth by the writer and in reaching beyond the page to establish relationships with previous knowledge and experience.[3]

For purposes of this study, the best way to define comprehension is to clarify what it is not; comprehension is not the score attained on either or both of the two tests in question. Scores on tests of reading merely sample a small number of reading skills which elicit an equally small number

---

[1]Charles C. Fries, Linguistics and Reading (New York: Holt, Rinehart and Winston, Inc., 1966), p. 93.

[2]Morris Finder, "Comprehension: An Analysis of the Task," Journal of Reading, XIII (December, 1969), 199-202, 237-240.

[3]Henry A. Bamman and others, World of Ideas: A Guide to Effective Reading (Berkeley Heights, N. J.: Field Educational Publications, Inc., 1970), p. 5.

of pupil behaviors. A judgment regarding comprehension is an inference made from these behaviors.[1]

To define comprehension as measured by any test it is necessary to determine what skills are being tested and the degree to which a subject manifests those skills becomes the basis for judging his level of comprehension within that framework. The New York State Minimum Competence Test in Reading is based on four skills as stipulated in the original test manual;[2] these are: recognition of central thought of the selection, questions on specific statements, inferences about the content of a selection, and derivation of word meaning from context.

A weighted analysis of the items on the Minimum Competence Test indicates that in order of priority, the emphasis on skills tested is generally as follows:

Type of Question	Percentage of Questions
Questions on specific statements	48%
Inferences concerning content	30%
Contextual clues to word meaning	14%
Main idea or central thought	8%

Since the Reading Ability Test was designed to be congruent with this pattern, it stresses these skills in a similar order of priority; the study does not maintain that these are the best set of skills to be tested nor that they are

---

[1]Robert L. Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education (New York: John Wiley and Sons, Inc., 1961), p. 161.

[2]New York State Education Department, pp. 6-7.

ordered according to the best priority. However, inferences made about how urban children respond to questions on the two tests constitute reading comprehension as it is considered in this experiment.

Minimum competence in reading. The New York State Minimum Competence Test in Reading equates the ability to apply the four skills just designated with the notion of minimum competence in reading. This study would be negligent were it to tacitly subscribe to that equation. The term, minimum competence, was rejected as a suitable descriptor in the title of the new, congruent test; this was because minimal competence in reading is both relative and subjective and cannot adequately be measured by forty questions which measure four skills.

Minimum competence in reading implies a facet of literacy; namely, the ability to decode symbols. Rogers points out the relativity of functional literacy in the following way:

> The exact nature of literacy varies with different roles. As the requirements for a given position in society vary, so does the degree of literacy required to adequately cope with the position. For example, the level of literacy required for functioning as a subsistence farmer is very different from that required for a university professor.[1]

The subjective nature of minimum competence is also

---

[1]Everett M. Rogers, Modernization Among Peasants: The Impact of Communication (New York: Holt, Rinehart and Winston, Inc., 1969), p. 205.

readily apparent. Reducing reading competence to its simplest terms, Pittenger[1] explains that no one reads to read; one reads in order to . . . ; therefore, reading always involves purpose of some kind. That purpose may stem from an exterior coercion, such as, the need to interact with one's culture; or, purpose may come from a personal need, such as, the need to think or feel through the printed word.

The subjective and relative nature of minimal reading compentency does not imply that it would be impossible to construct an adequate test of minimal reading ability. However, such a test should be based on those external cultural coercions which demand that an individual decode symbols if he is to live safely and happily in his environment. Gates and Jennings[2] refer to "vital reading" tasks and these would seem to constitute the true basis for a test of minimal reading ability. Such tasks as reading job applications, menus, directions, social security forms, driver's manuals, applications, guarantees, insurance forms, newspapers and other real-life reading activities logically constitute a body of realistic reading processes on which a valid measure of minimal reading competence could be constructed. Such a base, no

---

[1]Owen Pittenger, "Psychological Consideration in Identifying Disadvantaged Readers" (paper read at a Reading Symposium, August 12, 1969, Oswego University, New York).

[2]Arthur I. Gates and Frank G. Jennings, "The Role of Motivation," Development In and Through Reading, Sixth Yearbook of the National Society for the Study of Education (Chicago: University of Chicago, 1961), pp. 121-122.

doubt, would be considered too narrow for those who claim that to be able to read vital matter and nothing more is to be semi-literate. Members of the Reading Reform Foundation[1] subscribe to this theory and feel that a high school graduate should be able to cope successfully with such reading tasks as those presented by the New York Times, the Christian Science Monitor, Time, Newsweek, etc. No educator could argue with this goal; however, it seems to exceed the notion of minimal reading ability which implies the least possible skill in reading that one needs in order to cope adequately with the necessary reading tasks imposed by his environment.

The point of this lengthy investigation into the term, minimum competence in reading, is to demonstrate that it is an unsuitable label for either of the tests in this study. Test results will hardly verify that a child who scores above the magic number of twenty-six is qualified to handle functional, real-life reading tasks. It would seem that either the descriptor, minimum competence, should be dropped from the title of the New York State Test, or the test itself should be designed to fit that descriptor; the latter is obviously preferable.

Both the New York State Minimum Competence Test in Reading and the Reading Ability Test sample a small number of behaviors from which some indication of reading comprehension

---

[1]Watson Washburn and Robert C. Price (eds.), Proceedings of the Eighth Annual Reading Reform Foundation Conference (SanFrancisco: Reading Reform Foundation, 1969), p. 15.

may be inferred but as measures of minimal competency as it relates to functional literacy, both tests fall short. Competency in this study is construed as a much broader term implying the ability to handle life's necessary reading tasks.

Content validity. Defined within a narrow perspective, content validity concerns itself with the degree of match which may be found to exist between test content and related course content. It may, however, be defined within broader terms and it is this extended definition which is most suitable for this study. Thorndike clarifies this extension in the following manner.

> Since the analysis (course content and test content) is essentially a rational and judgemental one, this is sometimes spoken of as rational or logical validity. Since the analysis is largely in terms of the content of the test, the term content validity is also sometimes used. However, we should not think of content too narrowly, because we may be interested in process as much as simple content.[1]

This study concerns itself with what an urban child does, how he behaves toward two types of content--one relevant and one alien. Process and content are of equally great concern, therefore, a broad definition emphasizing process is adopted for this study.

Basic Assumptions

Characteristics of children in the study. Because of

---

[1] Thorndike and Hagen, p. 162.

discrepancies noted in the Introduction to this study, namely that between statistical and empirical evidence of reading ability, the basic assumption of the study was formulated. That assumption was that a number of children in urban areas possess a greater degree of reading ability than is manifested in typical reading test results. Although the assumption was made concerning children who generally score low on standardized tests of reading, the assumption may be equally true of children who are high scorers. Even capable readers may score consistently higher on more relevant reading materials. Thus, a mean score increase found in this kind of comparison may reflect better reading achievement of both poor and good readers, as well, on relevant reading material. An investigation of the extreme groups may reflect such findings.

It is further assumed that a child's basic interests, values, and the experiential base on which these were formed have an immediate effect on his reading achievement. This would be true, however, only if, in addition to having the necessary experiential base, the child has the innate intellectual ability to profit from his experience. For this reason, mentally deficient and brain injured children have not been included in the study; it was assumed that for them, little, if any, difference would be found in their reaction to two distinct kinds of reading material.

Furthermore, it was assumed that it is possible to identify specific interests and values which would be common to many urban children.

Another assumption which is crucial to this study is that current emotional and sociological factors among urban youth may be deflecting reading achievement scores. For example, racial tension found among children in urban schools today may be affecting testing results. To cite an example, the question of what effect the first paragraph of the New York State Minimum Competence Test in Reading has on a black pupil is a valid one. The paragraph deals with albinos and such phrases as "white in color," "white lions," and "milky-white skin," could actually spark a negative reaction in black children. They could understandably have a poor response to the test as a whole. The fact of their being measured by a tool designed for and by a white society could interfere with a true measurement of their reading ability. Such a conclusion may have been inconceivable several years ago but is feasible today in light of the racial problems so evident in our society. Black educators and psychologists are particularly sensitive to problems of the nature; Bibby demonstrates this sensitivity in the following statement:

> The white man's pride of colour is understandable as a reflection of his several centuries' technological superiority, but it has become associated with the idea that 'whiteness' itself is in some way superior to brownness or yellowness or blackness. This complex emotional content of the word 'white' would become apparent to children of European origin, if for a period, they and their teachers deliberately used the rather more accurate epithet 'pink' in classroom discussion of ethnic differences.[1]

---

[1]Cyril Bibby, Race Prejudice, and Education (New York: Frederick A. Praeger, 1960), p. 63.

Black psychologists throughout the country are acutely aware of the affective nature of the black child in particular, and his psychological response to testing situations. The Association of Black Psychologists[1] in SanFrancisco, for example, conducted a successful drive to halt ability and intelligence testing until equitable means of ability assessment are developed. Black doctoral students[2] at the University of Michigan have voiced the need for research which will reflect the needs of black children.

This study is an effort, limited though it may be, to respond to that need. Although it was designed for both black and white children in an urban area, approximately one half of the children in the sampling were black and a comparison of their test responses appears to be in accord with the exhortations of the black psychology-oriented groups just described.

Nature of tests in the study. The New York State Minimum Competence Test in Reading may be an inaccurate measure of the reading ability of urban children if it penalizes the child whose experiences have been limited by economic and/or sociological factors throughout his development. The concepts on the New York State Minimum Competence Test in Read-

---

[1]"Demands and Results," Behavior Today: The Human Sciences Newsletter, July 6, 1970, p. 1.

[2]News item in Behavior Today: The Human Sciences Newsletter, August 31, 1970, p. 2.

ing are construed as alien to the interests, values, and experiences of many urban children. In addition, the nature of the material itself is such that it sparks little personal concern or interest so that poorly motivated children find little reason in the test itself to read with interest or enthusiasm.

Contrary to the Minimum Competence Test, the Reading Ability Test was designed to increase the inherent interest factor thereby improving conditions for increased achievement. The interest areas, derived from urban children themselves, are considered to have a stronger motivational force while, at the same time, maintaining an equal degree of reading difficulty.

Lastly, it was assumed that it would be possible to construct a measure which could be congruent with another measure by the substitution of a different type of content within the same structure. No evidence could be found of any other test which purports to do the same thing; therefore, the Reading Ability Test itself is assumptive.

SIGNIFICANCE OF THE STUDY

Affecting Test Change

It is anticipated that this experiment may provide findings which could be incorporated into a revision of the New York State Minimum Competence Test in Reading. If urban children are more apt to score minimum competence on relevant reading material that fact cannot be ignored by test makers.

This experiment is concerned with the effect of content on reading achievement. A colleague of this writer, Mr. Ralph Ware, Director of the Curriculum Laboratory of the State Education Department, is making a study of how the method of questioning children on a test of reading affects the degree of their achievement. Ware is developing "an exemplary test for assessing the achievement of a particular set of comprehension skills in reading for a given population."[1] His population sample is the same as that used in this study; it is anticipated that an examination of the results of both studies will contribute much to an understanding of how and why urban children respond as they do on tests of reading ability.

Affecting Attitude Changes

What may be the most important outcome of this investigation is the possibility of raising the expectancy level of teachers, children, and even parents as a result of the findings. Unfortunately, after eight years of schooling, if children have been continuously under the impression that they are poor readers, this poor self-image becomes so ingrained that many children enter the ninth grade convinced that they cannot read and, as a consequence, see no point in trying. If through the Reading Ability Test it could be demonstrated

---

[1]Ralph H. Ware, "The Development of a Criterion-Related Reading Comprehension Test for Exposition and Argument" (unpublished Dissertation Proposal, State University of New York at Albany, New York, 1971), p. 1.

to these children that they can and do read well on certain kinds of material, significant attitude changes could result.

Such a reaction occurred two years ago in an Albany school where a trial-run of the two tests was made during the early stages of this experiment. After receiving results which indicated that children do read better on relevant material, the principal requested that the same testing be done in another grade because both teachers and pupils reacted so favorably to the encouraging results. To multiply this kind of reaction would be to make a significant impact on the teaching-learning process.

The first trial-run yielded an interesting anecdotal remark from one of the testers; she observed a marked difference in class attention during the two tests. Groups taking the New York State Minimum Competence Test in Reading were restless and the room was noisy with pencil-dropping, coughing, and general movement. During the Reading Ability Test however, this restlessness was absent and pupils appeared much more involved in what they were doing.

Initiating Similar Research

Another point of significance of this study concerns itself with other tests at other grade levels. Should it be demonstrated that the nature of reading content does have a positive effect on level of achievement, the experiment should be replicated in other places to determine if what is true of a group of children in Albany is also true of children in other urban areas as well. Furthermore, if a difference occurs

in grade nine, the New York State Pupil Evaluation of Performance Tests given in grades three and six should also be examined by using a similar method. In addition, commercial tests of reading could be scrutinized in like manner. One who has the time and the patience to write a parody-type test could follow the pattern of investigation established in this study. It is also possible for the Reading Ability Test to be given, as it stands, at other high school levels; many related studies which cut across geographical areas and grade levels could add to the significance of this one.

Bearing on Current Problems

This experiment may support the efforts of psychologists, educators and sociologists who raise objection to cultural bias in reading and intelligence testing. The Reading Ability Test was designed not as a culture-free nor necessarily culture-fair test, but as one which is culturally-relevant to a designated population. The test is not presented as an answer in itself to the problem of valid testing of urban children; it may, at best, shed some light on the problem of why urban children tend to score so poorly on standardized test measures of reading. Only if repeated trials are made with this test and other congruent tests developed, will the body of proof become substantial enough to bring about changes in state and commercially prepared measures. Testing one small group of children in one small urban area will hardly solve a national problem; however, it may stimulate a wave of research which can have widespread effects.

Chapter 2

RELATED LITERATURE

CURRENT CONCERNS

Standardized Testing

The subject of the suitability of standardized testing practices in urban areas is of immediate national concern; in the words of Dr. Roger T. Lannon:

> The discussion (concerning inner-city testing) has moved off the pages of educational and psychological journals onto the pages of mass media. Its forum has moved from the classroom and the psychological laboratory to City Hall and the courtroom.[1]

The question of the relevancy of test items on standardized measures of reading and intelligence is of equally great concern; the awakened interest in the humanistic aspects of curriculum contains within it the derived concern for the humanistic aspects of the testing process. If it is no longer sufficient to design curriculum with cognitive ends as the sole objectives of education, it is no longer sufficient to design tests of any kind without consideration for the affective nature of the child.

Test suitability in itself is not the major cause of

---

[1]Roger T. Lannon, "Testing: The Question of Bias," Evaluation in the Inner City, ed. Dr. Thomas J. Fitzgibbon (New York: Harcourt, Brace and World, Inc., 1970), p. 43.

mounting national concern for the testing process; the underlying concern is for the child who is judged and categorized on the basis of his performance on standardized tests. Testing for selection raises doubts as to social and moral justification as pointed out by Skager.[1] He refers to the collective conscience of the country which reminds us that test procedures which foster selection policies may, in fact, violate equality of opportunity. An awareness of the way in which institutions influence human development is dawning on both professional educators, psychologists, and parents throughout the country. Leacock[2] points out that knowledge of the effect of primary institutions, such as the school, on individual behavior is still in its early stages and although inroads have been made, much remains to be learned about so complex an interaction.

Miller delineates the testing dilemma and pinpoints the central issue in this way:

> Since individual differences are a function of many factors, how does one insure that in any particular test relevant factors are being measured and irrelevant ones excluded? . . . Essentially, the issue involves values; that is why it is so persistent and so resistent to the impact of research.[3]

---

[1]Rodney W. Skager, "Objective Based Evaluation: Macro-Evaluation," Evaluation Comment, II (1970), p. 7.

[2]Eleanor Leacock, Teaching and Learning in City Schools: A Comparative Study (New York: Basic Books, Inc., 1969), p. xx.

[3]Harry Miller, Education for the Disadvantaged (New York: The Free Press, 1967), p. 87.

The same dilemma exists in this study. There is no guarantee that relevant factors in terms of minimal reading competence are being measured and irrelevant ones discarded; however, some light may be shed on the affective response of the urban child in relation to two separate bodies of content. Miller implies the feasibility of such a study as this in his statement that it may be possible "to eliminate irrelevant cultural differences at the point of test construction."[1] Such attempts, he points out, are not likely to be entirely successful but may, at least, identify potential problems for test users and provide them with bases for clinical judgments. This study proposes no more.

Miller cites a study done by Cardall and Coffman which was designed to examine various factors contributing to differences in test performance among different groups. Findings of this study indicate that there is still no answer to the question of differentiating between balance and bias in a test. The factors creating significant interactions were not readily apparent from an examination of item content. "The factors producing changes in the relative difficulty of questions from group to group appeared to be the result of complex factors such as those producing individual differences within groups."[2]

Current concern for the validity of measurement tools

---

[1]Miller, p. 87.

[2]Miller, p. 89.

and the proper utilization of data which those tools yield, extends beyond the educational research laboratory and the schools themselves into urban communities. Parents of urban children demonstrate that they are sensitive to the problems involved in testing children fairly and utilizing data properly. Mrs. Novella Williams, a representative of parents in urban communities, voiced this sensitivity in her speech to a group of professionals in the field of testing:

> Your tests are irrelevant, foreign in many instances, because you're testing the inner city child in something he doesn't know anything about. . . . You cannot test any group of people on the face of this earth unless you get yourself some people from the life style of the people who you're testing and incorporate their ideas and their thinking into your own. . . . mentally we're in two different societies.[1]

The disparity which is often found between the cultural values which emerge from test items and the lived-cultural-values of many urban children results in what Dr. Charles Smith from the Rockefeller Foundation refers to as colored test results.[2] Tests of reading, in particular, which fall outside the range of a child's experiential background test not only his ability to decode words but, in addition, his ability to cope with strange, unfamiliar, and often confusing

---

[1]Novella Williams, "Community Concern About Testing: A Parent's Point of View," Evaluation in the Inner City, ed. Dr. Thomas J. Fitzgibbon (New York: Harcourt, Brace and World, Inc., 1970), p. 43.

[2]Dr. Charles H. Smith, "Crisis in Urban Education: Schools on the Offensive," ibid., p. 17.

concepts. The great danger and inherent injustice of specious results of such testing is that it is possible for a child to become locked into a low ability grouping because of these results and his true potential may never find release. In a society based on democratic principles this becomes an unpardonable captivity.

Recent studies done in the area of self-fulfilling prophecies reveal even more clearly the need to scrutinize both the design of standardized tests and the way in which results are utilized. Rosenthal[1] conducted an experiment with children in a low socio-economic neighborhood on the West Coast. Random groups were labelled as high scorers for "intellectual blooming" and described as such to the teachers in the experiment. Actually, no real difference existed between the intellectual bloomers and the remainder of the classes; the difference was in the minds of their teachers. At the end of the year the total gain in IQ of the experimental group was four points more on the average than that of the control group. The year-end survey of teachers indicated that affective changes occurred as well; teachers, for example, found children in the experimental group "slightly more appealing, better adjusted, and more affectionate, and with less need for social approval."[2]

---

[1]Robert Rosenthal, "Self-Fulfilling Prophecy," Psychology Today, II (September, 1968), 46-51.

[2]Rosenthal, p. 47.

The reverse of these results is apt to occur when a group of teachers find themselves faced with classes labelled as "low achievers," "poor readers," or "educationally disadvantaged" as a result of standardized testing.

Testing and Curriculum

Another immediate concern related to testing is the effect which testing has on the curriculum itself. Ebel[1] raises the point that any tool of measurement should closely parallel what pupils have learned. When the school uses external tests, such as published standardized or statewide measurements, the parallel between curricular teaching objectives and measurement of the attainment of those objectives is hardly ideal. The measurement process can be inaccurate and misleading in terms of what the child has actually learned through the prescribed curriculum. This fact is particularly true of the New York State Minimum Competence Test in Reading, the major consideration of this study. If, for example, the specific study of albinos, the British Isles, and pioneer problems were formal studies in the New York State Social Studies Curriculum, and if all teachers of social studies spent adequate time teaching these concepts, then the first three paragraphs of the New York State Minimum Competence Test in Reading would be completely fair and valid measurements of both reading ability and social studies

---

[1]Robert L. Ebel, "The Curriculum and Achievement Testing," Educational Technology, V (May, 1970), 22.

achievement for urban children. Since, however, these conjectures cannot be verified, the test items are questionable in terms of what children could reasonably have been expected to learn.

Furthermore, Ebel shows that external testing leads to several abuses which disturb curriculum specialists and administrators. He summarizes these as follows:

> . . . teachers are apt to 'teach to' a test
>
> . . . curriculum control lies in the hands of test makers
>
> . . . testing fosters uniformity of curricula
>
> . . . prepping for tests fosters undesirable teaching methods
>
> . . . written tests measure cognitive competence; other equally important areas (social, behavioral, affective) are neglected[1]

Norm-referenced Versus Criterion-referenced Tests

A further point of consideration pertinent to testing issues of national concern is that of the advisability of typical norm-referenced measures. Glaser's[2] descrîption of norm-referenced testing indicates that such measures show how an individual performs in relation to his peers. They do not indicate what a student has learned as an individual nor do they yield information as to his peculiar strengths and weaknesses; they rather predict the likelihood of his success

---

[1]Ebel, p. 22.

[2]Robert Glaser, "Instructional Technology and the Measurement of Learning Outcomes: Some Questions," American Psychologist, XVIII (August, 1963), 519-521.

as compared to that of his peers. Popham and Husek,[1] on the other hand, define a criterion-referenced test as one which measures performance in relation to specified criteria. Such a test provides no normative base for comparisons nor does it interpret what attainment of objectives means; however, it provides a more suitable measure of an education program.

Klein proposes that it is possible to combine the two distinct forms of measurement into one criterion-normative based test. He delineates the steps which would be necessary in this combination:

1. Specification of objectives
2. Development of test items for each objective
3. Development of test items to measure related objectives
4. Provision of score and score interpretation for each objective[2]

Such a procedure might serve as an ideal base on which to reconstruct the New York State Test of Minimum Competence in Reading.

Testing and Financing

Although the New York State Test of Minimum Competence in Reading is considered a rough estimate of reading ability

---

[1]James W. Popham and R. R. Husek, "Implications of Criterion-Referenced Measurement," Journal of Educational Measurement, VI (Spring, 1969), 1-9.

[2]Stephen Klein, "Evaluating Tests in Terms of the Information They Provide," Evaluation Comment, II (June, 1970), 1.

it serves normative purposes. The State Education Department computes the total picture of children's reading ability by listing percentages of children reading below so-called competence in reading. Cities in New York State are compared percentage-wise and individual districts are notified as to how they rank on a state-wide scale. Klein delineates the dangers involved in use of test results:

> Thus, test results wield enormous power in educational decisions that determine the kinds of educational programs a student receives and the level and direction of his educational career. These, in turn, influence greatly his place in society.[1]

The implications of the power of test results are particularly formidable today when increasing federal and state supplementary funds are being directed into education. For example, test results in the area of reading throughout New York State have played a crucial role in direction of the bulk of ESEA Title I monies into the improvement of reading programs at every level of elementary and secondary education. If, by any chance, measuring devices have been invalid or inappropriate and if the picture of reading achievement state-wide is not as drastic as originally thought, other equally pressing needs of education may have been neglected. This, obviously, is a hypothetical situation and would not be true unless research intesting could prove the contrary; however, it does illustrate the effect which test results exert on educational projects. To

---

[1]Klein, p. 1.

multiply the effects of this hypothetical case nationwide illustrates the tremendous influence of testing on educational programs. More than ever, measures of reading achievement, in particular, must precisely reflect actual abilities if funds are to continue to be allocated properly.

The national assumption, based on standardized, normative measures is that reading deficiency is the primary problem in education. This assumption, based on standardized, normative measures is that reading deficiency is a critical problem nationwide. Such facts as the following, compiled by the United States Office of Education, underly this assumption:

> 7 million public school pupils in the United States require special reading instruction.
>
> 40% of secondary school pupils require special instruction in reading.[1]

The question which arises in light of these statistics is that posited by the Introduction of this study. How can the profiles across the country be so bleak when there is so much empirical evidence in classrooms which attests to optimism? Children have never read more widely than they do today. Publications of all types pour into schools and are fully utilized. With all that is available, teachers never have enough reading material for their pupils in every subject area of the curriculum. The consumption of paperbacks is at an all time high. Language elements in elementary

---

[1]National Center for Educational Statistics, Advance Statistics for Management, V (April, 1970), 1.

grades are introduced and mastered sooner than they were fifteen, ten, and even five years ago. Yet despite the tangible, optimistic evidence of this nature, test profiles remain grim.

If these statistics are to remain the basis for allocating additional state and federal funds, the tools by which they are gathered deserve careful examination so as to insure that test profiles accurately reflect actual reading abilities.

PROFESSIONAL OPPOSITION

The greatest repercussions from this study will most likely issue from test makers who will argue that there is no basis, no precedent, for the construction of a congruent test. The idea of substituting familiar content within a similar test structure will most likely be attacked on the grounds that the introduction of new content creates a completely dissimilar test. Parallel test forms are a recognized and approved method of drawing comparisons because they sample the same "content universe" as described by Helmstadter.[1] The Reading Ability Test changes that universe but does so within the same test framework and at the same level of difficulty.

There will be need for replication of this experiment at other times in other places to determine if credence may be extended beyond the confines of the City of Albany. There

---

[1]G. C. Helmstadter, Principles of Psychological Measurement (New York: Appleton-Century-Crofts, 1964), p. 66.

is also need for construction of other congruent tests at other levels before a solid base of argument may be presented to test makers who reject the notion of congruency in testing.

A second, but equally strenuous objection, will probably be raised by those who maintain that reading and intelligence tests currently in use do exactly what they are designed to do and that no revision or scrutinization of tests in themselves is necessary. Cohen,[1] for example, defends the use and validity of standardized tests and states that "it does not seem wise to blame the tool (test)" for indicating what is the obvious fact; namely, that the disadvantaged, as a group, are poor achievers in schools which are based on middle class values. Cohen sees the pedagogical process itself as the principal cause for poor test results.

Subscribers to Jensen's[2] philosophy will also be quick to censure this experiment on the basis of the hypothesis that a segment of the population will never be able to measure up to standardized norms because of intellectual deviations or deficiencies. Jensen was not the first to raise the question; he had been preceded by many people who felt that intellectual deficiency explained the inability of a number of

---

[1] S. Alan Cohen, Teach Them All to Read: Theory, Methods, and Materials for Teaching the Disadvantaged (New York: Random House, 1969), p. 56.

[2] A. R. Jensen, "How Much Can We Boost I.Q. and Scholastic Achievement?" Harvard Educational Review XXXIX (1969), 1-123.

children to cope with the reading problem. In 1963, for example, Gans made the following statement:

> The problem, bluntly stated, is the failure of the secondary school to face the fact that upwards of thirty per cent of our wholesome youngsters are not born with the intellectual potential either to acquire a certain degree of skill in reading or to learn through a method described by an able youth as "throwing the book at kids."[1]

Children who have been thoroughly diagnosed and identified as brain injured, mentally retarded, or seriously emotionally disturbed have not been incorporated in this study. For content to have an impact on comprehension, basic intellectual powers must be functioning. Children, however, who are considered slow learners, retarded readers, and mildly disturbed emotionally, have been included and are of major concern. There is great danger of their being penalized unfairly by the nature of reading matter which is used to categorize them as non-competent readers. Should the findings of this study prove significant, there will be a need to investigate if the alien nature of concepts on intelligence tests as well create a similar distortion. There is a distinct possibility that some children are currently labelled as non-competent in reading when they can, indeed, decode words and respond correctly to familiar content but find it impossible to deal with totally unfamiliar concepts. The same error may exist in cases of intelligence testing.

This study does not maintain that all children,

---

[1]Roma Gans, Common Sense in Teaching Reading (New York: The Bobbs Merrill Company, 1963), p. 395.

regardless of intelligence, will respond favorably to relevant reading matter. It does, however, posit that the nature of reading matter presented on tests of reading could be distorting true levels of ability among some children.

A third objection to this type of study will be that changing the content of a test in order to take into consideration cultural familiarity will merely serve to vitiate the test instrument. Horricks notes the severe analytic and methodological problems involved in taking "fully into consideration the more molar cultural aspects of capacity to perform"[1] on intelligence tests. He cites the example of the social scientists at the University of Chicago who sought to develop a culture-fair rather than culture-free test. The Davis-Eells Games grew out of this attempt; however, research revealed adverse findings and low predicitive validity of the Games. Horricks[2] pointed out that despite this failure to develop a test free of socio-economic bias, there are interesting research possibilities which deserve further exploration. This study attempts to further that exploration in attempting to construct a test of reading which may be a fairer measure of reading ability for the urban child. Although the Reading Ability Test alters content in light of cultural variations, the test is not vitiated since its read-

---

[1]John E. Horricks, Assessment of Behavior (Columbus, Ohio: Charles E. Merrill Books, Inc., 1964), p. 283.

[2]Horricks, p. 285.

ability level is matched with that of the New York State Minimum Competence Test in Reading. The two tests are equally difficult but content differentiated.

Equating two tests by means of a readability formula will bring censure from those who condemn readability formulas as predictors of difficulty. Dawkins,[1] for example, has attacked the Dale-Chall Formula of Readability as (1) not considering words in terms of their context, syntactical position or grammatical structure and (2) not evaluating structure of the paragraph, organization of the content or difficulty of the subject matter. He doubts any correlation exists between the formula and actual reading difficulty.

Spache,[2] on the other hand, defends the use of readability formulas as valid and justified for specific purposes. Although Dawkins' arguments are feasible, there is no statistical way other than the use of a formula to estimate the difficulty of two tests. Subjective judgment alone is hardly an improvement over a readability formula. Therefore, in order to establish some base of comparison the Dale-Chall Formula was adopted; the formula provides a working basis for comparison despite its alleged weaknesses.

A final objection to be considered here will be that posited by test constructors who maintain that in order to

---

[1]John Dawkins, "A Reconsideration of the Dale-Chall Formula," Elementary English XXXIII (December, 1956), 515-519.

[2]George D. Spache, Good Reading for Poor Readers (Champaign, Ill.: Garrard Publishing Company, 1966), p. 33.

arrive at a true measure of pure reading achievement, the content of a test must abstract from a reader's experience. If the reader functions within the range of his experience, proponents of this school of thought feel that the measure is not a true indication of reading ability. In other words, if a reader can handle alien content successfully, he has the reading skills necessary to do so and his score is in no way influenced by the familiarity of the material since material presented is purposely chosen as irrelevant in relation to the test population.

This is, without doubt, a solid argument and has definite implications for the construction of test items; however, when the subject of the testing is an urban child and the goal of that testing is establishing whether that child has minimal ability to read, selecting content which abstracts from his experience may penalize him unnecessarily. Many children in urban areas grow up deficient in the extent of their conceptual development due to the physical and/or sociological limitations on their experience. Nevertheless, many of them can and do read well within the scope of that experience. To test such a child with material written beyond such a framework is to test him on two counts: namely, to test his ability to decode words while at the same time testing his ability to deal with unfamiliar concepts.

Stauffer,[1] Warner,[2] and Lee and VanAllen[3] attest to the singular success of the language-experience approach to reading which creates the necessary link between what a child says and how his speech appears to him when reflected in print. The language-experience philosophy has been adapted to some degree in the Reading Ability Test in that reading material has actually emerged from the conversation of urban children.

PROFESSIONAL SUPPORT

Much support for the kind of investigation proposed by this study was implied in the first section of this chapter which dealt with current concerns about testing found throughout the country today. However, more substantial and specific support is in order.

Concept familiarity as it relates to comprehension forms the basis of this study. Therefore, it is appropriate to look into the literature which deals with the concept-

---

[1]Russell Stauffer, The Language-Experience Approach to the Teaching of Reading (New York: Harper and Row, Publishers, 1970), pp. 268-285.

[2]Sylvia Ashton Warner, The Teacher (New York: Simon and Schuster, 1963), pp. 27-100.

[3]Doris M. Lee and Roach VanAllen, Learning to Read Through Experience (New York: Appleton-Century-Crofts, 1963), pp. 1-2.

reading relationship. Both Piaget[1] and Brodbeck[2] point out the difficulty involved in forming concepts through speech; Piaget feels that concepts cannot be taught verbally but must include some type of activity which relates to them. Brodbeck illustrates the difficulty of forming concepts verbally with the following example. If one has never seen a gnu and must rely on the verbal description of a mammal with an oxlike head, short mane, downward-curved horns and a long tail, he will have difficulty in piecing these concepts together despite the fact that he may be thoroughly familiar with each separate idea. If a gnu per se is not registered in his sight-memory-bank, his concept, based solely on verbal description, is likely to be a faulty one. This study maintains that children, particularly children in urban areas, have no less difficulty with concepts of albinos, the British Isles, pioneer problems, etc.

It is obvious that concept formation is a complex process and a test writer may be tempted to many false assumptions concerning the nature of universal content. Adjusting test content to suit the concept familiarity of children in different localities has been tried before. Storey[3] recog-

---

[1]Elizabeth Hall, "A Conversation With Jean Piaget," Psychology Today, III (May, 1970), 30.

[2]May Brodbeck, "Logic and Scientific Method in Research on Teaching," Handbook of Research on Teaching, ed. N. L. Gage (Chicago: Rand McNally and Company, 1963),p. 48.

[3]Arthur George Storey, "A Reading Comprehension Test for Junior High School Pupils in Alberta," (unpublished Doctor's dissertation, Stanford University, Canada, 1956), pp. 1-61.

nized the unsuitability of using commercial reading tests produced in the United States to measure the reading comprehension skills of Canadian children; he therefore designed a test to be used locally in the Alberta Junior High Schools. Content of his test centered around such topics as: egg hunting, forest life, animal foods, Canadian democracy and the English language. His test was a replacement for the California Reading Test which contained such items as: the annual output of oranges in the United States, geography of Switzerland, the first bridges, diagonal lines, the terrain of Italy, and the Dewey decimal system. Although Storey states that his intent was to design a test to measure reading ability on the basis of the curriculum in the Alberta schools, he automatically brought the test material within the actual experiential range of the children being tested.[1] His purpose was not the construction of a congruent test, nevertheless, he shares with this writer the conviction that reading skills must be tested within a logical and familiar framework rather than on an alien basis.

Cohen,[2] Smith,[3] and Carroll[4] consider various aspects of concept formation as it relates to the reading process.

---

[1]Storey, p. 4.

[2]Cohen, pp. 51-52.

[3]Henry Lee Smith, English Morphophonics: Implications for the Teaching of Literacy, Monographs of the New York State English Council, No. 10 (New York: New York State English Council, 1968), p. 3.

[4]J. B. Carroll, Language and Thought (Englewood Cliffs, N. J.: Prentice-Hall, 1964), p. 104.

They are in agreement that reading consists of processing a symbol of a symbol. Smith states that

> Language, by its very nature, is a symbolization of the experience of a group of human beings who share a culture, and writing is a symbolization of language. Writing thus must be seen as symbolization of a symbolization.[1]

Cohen[2] points out that prior to reading, a child must possess the related oral-aural symbol of what he is to read and be able to associate the written symbol with the oral-aural symbol which grew out of his experiences. The Reading Ability Test was based on this interplay between written and experiential symbols and was designed to increase the degree to which this interplay takes place among urban children taking a reading test. Although a child may be able to decode a written symbol, if he lacks its conceptual counterpart, the symbol remains, for all practical purposes, useless. The New York State Test of Minimum Competence in Reading presents children with many concepts which may be alien to them. Concepts of pioneers, rennet, albinos, architecture, Puget Sound, the village blacksmith shop, etc., are likely to be completely lacking or insufficiently developed among city, block-oriented children and are, for that reason, considered to penalize an urban child in his attempt to crack a reading code.

This study does not deny that a great advantage in being able to read is that an individual may enlarge and

---

[1]Smith, p. 8.

[2]Cohen, p. 52.

enrich his experiences. According to Lerner,

> The very act of achieving distance and control over a formal language gives people access to the world of vicarious experience and trains them to use the complicated mechanism of empathy which is needed to cope with this world.[1]

This fact is completely accepted by this study, however, the study maintains that such laudable ends of reading do not belong in a test of minimal reading ability. Expansion of one's vicarious world is an essential facet of curriculum but the purpose of a test of minimum reading competence should not essentially include expansion of vicarious background; it should measure the extent to which one can cope with his everyday world and the reading functions which life in that world imposes.

In moving the New York State Minimum Competence Test in Reading from outside a probable range of interests and values of urban youth to within a more feasible experiential range, two major sources of information were utilized; these were urban children themselves and professional people who have developed guidelines for tapping the values and interests of children. Raths[2] provided the chief tools which were used by teachers in a summer communications course prior to the construction of the Reading Ability Test. Appendix H of this study illustrates the type of methods used by teachers

---

[1]Daniel Lerner, The Passing of Traditional Society: Modernizing the Middle East (New York: Free Press, 1958), p. 64.

[2]Louis E. Raths, Values and Teaching (Columbus, Ohio: Charles E. Merrill Publishing Company, 1966), pp. 3-258.

to identify pupil interests and values. Figure 1 is a copy of Carroll's "Meaning Space Chart" and demonstrates how concepts are rated on a value scale. Location on the axes shows the degree of activity or potency involved in each concept; small, as opposed to capital letters, indicate good concepts. In so far as possible, topics for the Reading Ability Test were selected in the active, strong, and good category with the intention of eliciting the greatest degree of inherent interest possible.

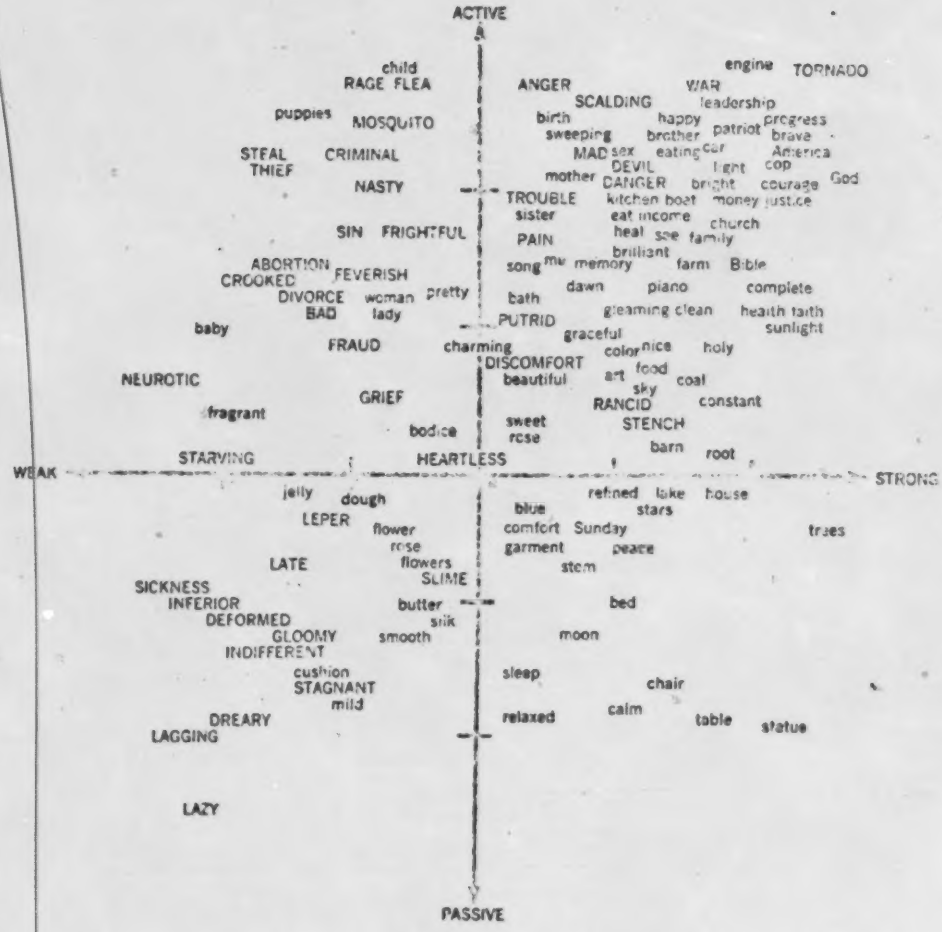
Further justification for incorporating familiar concepts as a means of creating a more equitable measure of reading was found in an experiment conducted by Smith and Hesse.[1] Recognizing that reading is always done in relation to existing background, they tried to discover what effect would be produced on comprehension by preparing children for new concepts in materials which they were about to read. Prior to an actual reading test situation, a three minute "cognitive organizer" was presented to the children in the study. This consisted of a tape presentation designed to familiarize children with the concepts about to be read and the sequence in which they would be encountered. The experiment showed that poor readers were definitely helped to read and comprehend better as a result of being prepared for what they were going

---

[1]Richard J. Smith and Karl D. Hesse, "The Effects of Pre-reading Assistance on the Comprehension of Good and Poor Readers," Research in the Teaching of English (Champaign, Ill.: National Council of the Teachers of English, 1969), III, No. 2, pp. 166-177.

Figure 1

Meaning-Space Concept Chart



Source: John B. Carroll, Language and Thought, (C) 1964. Reprinted by permission of Prentice-Hall, Inc., Englewood Cliffs, N. J.

to see in print. Better readers, on the other hand, showed no significant gains as a result of such preparation.

Ozman[1] conducted a study of the values transmitted through basal readers to elementary school children and maintains that educators are unaware of specifically what philosophies are being fostered among children through reading material of all kinds at all grade levels. He feels that there is an urgent need to examine all reading materials, tests included, since his findings show a lack of pluralistic values in the basal readers which he studied. Since testing situations are perceived as crucial experiences, it would seem that it is particularly important to examine what values are being transmitted and if they deserve the kind of automatic emphasis the test situation bestows on them. Such, however, would be the basis for another complete dissertation; the limitations of this study make such considerations impossible here; nevertheless, an awareness of their importance should not be overlooked.

Although much additional literature supporting the use of familiar concepts may be found, enough has been considered at this point to illustrate the rationale for the selection of the type of content chosen and the reason for its incorporation in the Reading Ability Test.

A second major area of support for this study comes

---

[1]Howard A. Ozman, "Value Implications in Children's Reading Material, The Reading Teacher, XXII (December, 1968), 250.

from the findings of linguists who are concerned with dialect differences found among urban children and how these influence the reading and learning processes. Erickson's[1] "Sounds of Society Project," conducted in Chicago studied the relationship between discussions among urban groups as compared with suburban groups. He discovered that both groups tended to shift back and forth between relatively restricted and relatively elaborated code depending on the context of the discussion. Restricted code refers to the abbreviated speech patterns which are used when shared content or similar experience is relatively high. Elaborated code, on the other hand, is more precise, extended speech which becomes necessary when the overlap of experience is low among people who are communicating. Although Erickson's[2] study based on Bernstein's Context Principle, related to speech rather than reading per se, the concept of shared context is similar to that proposed by this study. The idea of building congruent content into a test of reading ability was meant to increase the overlap of experience between the two communicators; namely, the child and the test. An interesting research possibility remains; the Reading Ability Test could be taken one step further and written according to restricted code which would match the spoken language of the urban child even more closely.

---

[1]Frederick D. Erickson, "'F'get You Honky!': A New Look at Black Dialect and the School," Elementary English, XXXXVI (April, 1969), 498.

[2]Erickson, p. 497.

Substantial support for this study comes from Baratz and Shuy[1] in their collection of articles which deal with dialect differences and reading relationships. The collection highlights the fact that overlooking divergence in speech patterns among urban children may underlie some basic causes of reading deficiency, particularly among black children in urban areas. Baratz emphasizes that ". . . children learn language in the context of their environment;"[2] this is a basic assumption among linguists and is applied to this study. If children learn language within a specified framework, it is reasonable that their reading ability be tested within, not beyond, that framework. Stewart[3] sees the ghetto child as dealing with a quasi-foreign language when that child is asked to deal with standard English language situations. Although the Reading Ability Test does not take dialect differences into consideration, it does, at least, reduce the foreign element by pulling content within a range likely to be understood by urban, ninth grade children.

Linguistic studies give indirect, nevertheless, important support for this study. A basic tenet proposed by linguists is that experience is an integral factor in language

---

[1]Joan C. Baratz and Roger W. Shuy, Teaching Black Children to Read (Washington: Center for Applied Linguistics, 1969), pp. 1-201.

[2]Baratz and Shuy, p. 96.

[3]William A. Stewart, "Foreign Language Teaching Methods in Quasi-Foreign Language Situations," Non-Standard Speech and the Teaching of English (Washington: Center for Applied Linguistics, 1964), pp. 1-15.

development. This tenet is the essential basis for this study. Just as the linguist recognizes an orderly, patterned language among ghetto children and bases teaching of language skills on that realistic basis, this experiment recognizes a distinct conceptual development among urban children and proposes the construction of a test of reading ability on that designated basis.

Another area of support which underlies this study is human motivation; basic to this study is the belief that the New York State Minimum Competence Test in Reading fosters little interest in an urban child. Many children in urban areas have little motivation to do well on tests of any kind and the slight motivation which does exist sometimes disappears completely when material presented on a test has no bearing on the child's immediate life concerns. Teachers attest to the fact that these children can and do read material which they find to be relevant. For example, a teacher of reading in a ghetto high school in Albany recently reported the amazement she felt when a class of children reading a new series of high-interest books failed to hear or respond to the bell which signaled the end of class. The Reading Ability Test seeks to capitalize on this type of potential.

Whipple and Black[1] characterize the plight of many

---

[1]Gertrude Whipple and Millard Black, "Reading for Children Without, Our Disadvantaged Youth," IRA Reading Aids Series (Newark, Del.: International Reading Association, 1966), p. 3.

junior high school children as ingrained failure. Children at this level frequently feel that they are too far behind to catch up and repeated failure, particularly with the reading process, breeds an actual hatred of the printed page. They recommend the use of reading materials with which these children can identify and in which they can find some personal value. Smith[1] supports the idea that all literature presented to children fall within realistic readability levels in order to prevent the frustrating experience of trying to cope with what is too difficult which ends in complete discouragement.

Getzels,[2] Morgan and King[3] agree that motivation to do or to refrain from doing something is characterized by three factors; namely, (1) the drive toward or away from something, (2) the striving or behavior which leads to a goal and (3) the satisfaction of achieving the goal itself. In this study the goal of each test is ideally the achievement of a passing mark; the striving is the effort expended by children in reaching that goal. The problem is that among many urban children there is little, if any, reason to aim for such a goal; their motivation is more apt to be negative and

---

[1]Nila B. Smith, "Literature for Space-Age Children," Education Magazine, October, 1960, p. 5.

[2]Jacob W. Getzels, "The Nature of Reading Interests," Developing Permanent Interest in Reading (Chicago: University of Chicago, 1956), p. 42.

[3]Clifford T. Morgan and Richard A. King, Introduction to Psychology (New York: McGraw-Hill Book Company, 1966), p. 236.

in the opposite direction of the ideal goal. Coburn[1] reports that urban children have little rapport with typical heroes of high school literature and that their need for relevant reading experiences is great. Classroom behaviors illustrate that when relevant materials are incorporated, startling attitude and achievement changes take place. Another illustration of this came from Sister Mildred, a school librarian in a ghetto school in Albany. In an unsolicited report she demonstrated this key to stimulating motivation among urban children:

> One day the Instructional Material Center Director made available to me several copies of magazines which feature cars. She suggested that they might be of some use in our small library. The following day, as I tried to present a "well planned" reading lesson to a group of ten-year-olds, I was confronted with a real problem. The boys had spied the new magazines in the rack and had become so absorbed in scanning the pages of Hot Rod, Motor Trend, Car and Driver and others that I just couldn't attract their attention. The thought came to me, "Capitalize on this interest!" After allowing the children ten or fifteen minutes to browse through the magazines, we launched a project that lasted almost three months.

Eastman[2] feels that children "turn off" when literature presented to them as a standard of excellence in no way satisfies adolescent needs nor corresponds with their problems. He feels that literature should be like a banquet table which

---

[1] Wilbur Coburn, Research and Development: Advances in Education (Washington: U. S. Government Printing Office, 1968), p. 87.

[2] Arthur Eastman, "After Literacy," English for the Junior High School Years (Champaign, Ill.: National Council of the Teachers of English, 1969), p. 113.

has something for everyone and a child should be able to pick and choose from among everything which is available. This assures reaching the affective domain of the child.

Some children who score poorly on tests of reading are actually good readers who, for one reason or another, refuse to reveal this ability. Kohl cites an example on one of these cases in the following dialogue:

> One day during the first week Alice coyly proposed a bet.
>
> "Mr. Kohl, I bet I can read anything on your desk no matter what those cards of yours say."
>
> Her reading score was 3.4. I accepted and she went through all the books on my desk including a page of the novel I was reading on the way to school. I was perplexed and delighted.
>
> "How can you do that and still have a three point four reading score?"
>
> "I wouldn't read for those teachers. Listen ---"
>
> Alice picked up a book and stumbled through several paragraphs. She paused, stuttered, committed omissions and reversals i.e. read on a low third grade level. Then she looked at my astonished face and burst out laughing.[1]

Pittenger[2] cautions against assuming that children can't read because they score poorly on tests of reading. A number of children can indeed read but "won't" or "don't" for many subtle reasons. Years of irrelevant reading experiences,

---

[1]Herbert Kohl, 36 Children (New York: The New American Library, Inc., 1967), pp. 184-185.

[2]Owen Pittenger, "Psychological Considerations in Identifying Disadvantaged Readers," (paper read at a Reading Symposium, August 12, 1969, Oswego University, New York).

more readily described as "phoney" by urban children, may breed an indifference to any act of reading which does not relate to a real-life-purpose. A test of minimal ability in reading is apt to reflect a more accurate reading picture if items chosen for it relate immediately to real-life-reading-tasks. The teenage boy, faced with his written driver's test, is not about to "turn off" when reading for comprehension on the test. The teenage girl is not about to ignore or skip over directions on a dress pattern when these relate to something she is making for herself. Short of such purposeful reasons for reading, a number of urban children have little tolerance for applying their reading ability.

Schick[1] looks very specifically into the types of motivation children do have to read. He maintains that reading is both thinking and feeling and that to these two processes there is no limit, no common denominator and no closed doors. He sees the reading process as a complex interaction and reaction and attempts to delineate the motives that anyone has for reading as follows:

- Relaxation
- Entertainment
- Knowledge
- Solution to Problems
- Reward
- Appreciation
- Aesthetic Development
- Identification with characters, situations, and conflict
- Escape
- Personal Development
- Curiosity
- Emotional Involvement
- Status
- Compulsion
- Vicarious Experience
- Insight
- Search for Self
- Development of Judgment

---

[1]George B. Schick, A Guidebook for the Teaching of Reading (Chicago: Psychotechnics Press Inc., 1966), pp. 19-20.

Unfortunately, the only motive which many urban children have for scoring well on a test of reading is that of compulsion. The Reading Ability Test was designed to add impetus to that lowly motive by incorporating topics designed to relax, entertain, make curious, involve and create identification. Haggard[1] states that a high level of motivation is particularly important in improving the test performance of lower class children. Test makers must take this into account in designing measures of achievement for urban children. A reading test situation for children who have experienced years of failure and frustration should be a task which is not only tolerable but actually enjoyable. Reduction of tension through content material which makes a pupil comfortable and less threatened may do much to arrive at a truer picture of his capabilities than content which threatens and confuses him by its lack of relationship to his experiential range.

An appropriate summary of the Basic Correlates of Reading Achievement and Deficiency (Figure 2) was constructed by Dechant[2] and illustrates the framework within which the reading ability of children should be judged.

---

[1]E. A. Haggard, "Social Status and Intelligence: An Experimental Study of Certain Cultural Determinants of Measured Intelligence," Genetic Psychology Monographs, XLIX (1964), p. 148.

[2]Emerald Dechant, Diagnosis and Remediation of Reading Disability (New York: Parker Publishing Company, 1968), p. 57.

Figure 2

Basic Correlates of Reading Achievement and Deficiency[1]

Growth in Reading Depends Upon

Adequate Experiential Background	Adequate Motivation
Adequate Language Background	Adequate Instruction
Adequate Maturation	Adequate Work Habits
Adequate Intellectual Development	Adequate Social-Emotional Development

[1]Emerald Dechant, Diagnosis and Remediation of Reading Disability (C) 1968. Reprinted by permission of Parker Publishing Company, West Nyack, N. Y.

Guilford points out that the essential question concerning the validity of a test must be not merely, "Is it valid?" but, "Is it valid for what?"[1] In this study the question proposed was: "Is the New York State Minimum Competence Test in Reading valid for judging the reading ability of urban children, in particular?" It is felt that the test fails to consider many of the basic correlates of achievement outlined in Figure 2 and is, therefore, not valid. Should a significant level of confidence be reached in the analysis of data in this study, that criticism will have been supported.

The proceedings of the most recent National Reading Conference point to the necessity for studies of the type developed in this paper. Jenkinson refers to the gaps in basic knowledge in the area of the personal determinants of reading performance and recommends research into "The problem of determining whether differing written genres or type of content result in differential reading performance, both between and within individuals."[2]

Carver delineates four conceptually different variables in reading; namely, knowledge gained, amount comprehended, reading aptitude, and reading improvement. He

---

[1]J. P. Guilford, Fundamental Statistics in Psychology and Education (New York: McGraw-Hill Book Company, 1956), p. 461.

[2]Marion D. Jenkinson, "Information Gaps in Research in Reading Comprehension," Reading: Process and Pedagogy (Milwaukee: National Reading Conference, Inc., 1970), p. 187.

launches a critical attack on traditional tests of reading comprehension and asserts that because of the vague definition of reading comprehension traditional measurement devices in reading are "measuring something other than the variable they purport to measure."[1] He feels that there is an urgent need to define reading comprehension in order that precise measurement tools may be developed accordingly.

Kingston proposes a Hierarchy of Reading Behaviors which forms the model on which realistic measurements could be constructed. He maintains that "there is some type of comprehension involved in even the most simple word recognition act"[2] and structures a set of reading behaviors which begins with this simplest act of comprehension and builds toward sophisticated applied and creative reading skills.

A substantial body of support currently exists for studies into the relationship between reading content and reading comprehension. Many additional experiments of the type developed by this study will be necessary before the relationship is more fully comprehended.

---

[1]Ronald P. Carver, "What Is Reading Comprehension and How Should It Be Measured?" Reading: Process and Pedagogy, ibid., pp. 99-100.

[2]Albert J. Kingston, "A Hierarchy of Reading Behaviors," Reading: Process and Pedagogy, ibid., p. 328.

Chapter 3

DESIGN AND PROCEDURES

HYPOTHESIS

The apparent discrepancy discovered in Albany between the steadily improving pattern of classroom performance in reading and the disappointing pattern of poor reading achievement manifested through annual test results gave rise to this study. It was conjectured that there may be some factor in measuring devices which accounts for the fact that test results fail to reflect encouraging growth patterns. The factor identified by this study as possibly defecting actual reading scores is foreign conceptual reading matter. In order to test if the nature of concepts presented affected the level of reading achievement attained, a new test was written which substituted familiar concepts within the same structure of a typical measure of reading ability. A null hypothesis was adopted which stated that: there will be no significant difference between reading scores attained by a group of ninth grade, urban children on two congruent tests of reading which are equally difficult but content differentiated.

The level of significance to reject the hypothesis was set at the 5 per cent level. It was anticipated this level

could be reached, and the hypothesis rejected. If a significant difference occurred, it would be attributed to the utilization of more relevant reading matter. Other variables such as format, level of difficulty, sentence structure, flow of thought, and pattern of questioning were held constant but the type of content presented was purposely altered in order to test its impact on comprehension.

TEST MATERIALS

The two tests which form the basis of this study are the New York State Minimum Competence Test in Reading and the test written to be congruent with it, the Reading Ability Test. The development and reason for use of each will be discussed at this point. Specific similarities and differences of each tool will be considered also.

New York State Minimum Competence Test in Reading

The New York State Test was chosen for this study for two reasons; first, it was designed to provide a means of measuring the reading ability of "pupils at the lower end of the scale of reading achievement."[1] As such a measure, it sets the standard for minimum reading standards for high school graduation; thus it is a crucial measurement for a number of children in urban areas particularly. Secondly, no commercially prepared tests of reading ability are used con-

---

[1]New York State Education Department, Manual of Directions for the New York State Minimum Competence Test in Reading (Albany: New York State Education Department, 1969), p. 5.

sistently and systematically throughout New York State; choice of reading measures is a local determination with the exception of the New York State Pupil Evaluation of Performance Tests given in grades three and six and the New York State Minimum Competence Test in Reading given in grade nine. These tools are used to construct statewide pictures of reading achievement; any such tools which serve as the basis for generalizing about the reading ability of large numbers of children deserve careful scrutinization.

The New York State Minimum Competence Test in Reading was constructed on the basis of a screening test run in February, 1962. According to the Manual[1] 4,000 pupils in thirteen school districts were exposed to 351 preliminary questions through eight pre-test forms. This preliminary testing excluded New York City. Each question in the random sampling was answered by approximately 250 pupils. On the basis of an item analysis the discriminating power of each item was determined and from the original 351 questions, forty were selected for the final form which is the Minimum Competence Test as it appears in Appendix A.

Determination of the minimum competence point of 27 was made through a May, 1962 testing of 1,700 pupils in grade twelve. Of this random sample 97 per cent obtained scores of 26 or higher. Pupils included in the sample represented Long

---

[1]N. Y. S. Education Department, pp. 10-11.

Island and upstate New York schools. The first over-all testing was made in 1965 when virtually all ninth grade pupils in both public and non-public schools of New York were tested. Public school pupils numbered 227,920; of this number 83 per cent scored 26 or higher. Public and non-public pupils combined totalled 271,166 and 85 per cent of these achieved minimum competence in reading.

Since 1965, the Minimum Competence Test in Reading has been administered annually to ninth grade pupils across the State early in each school year. Results are tabulated and utilized in the construction of statewide and local comparison pictures of reading ability. These annual results continue to be a source of great concern to administrators in urban areas, in particular, since every year a large number of children fail to exhibit minimum competence in reading on this test.

Reading Ability Test

Annual disenchantment with results of the Minimum Competence Test in Reading in Albany, New York gave rise to the development of a new test of reading ability. This test was designed to investigate if the nature of reading matter presented to an urban child influenced the level of his achievement. As noted in the Introduction of this study, a marked discrepancy has been observed between classroom performance in reading and testing performance in reading. High-powered, enriched, and expanded reading programs developed through supplementary State and Federal funding have caused

observable positive changes in the ability of children to cope with the printed word. Supervisors and teachers attest to the fact that children, as a whole, read better today than they did six years ago. An input of thousands of dollars every year into additional reading materials still does not satisfy the demand for more books, magazines, paperbacks, and reading matter of every type. It is difficult to explain why reading scores are not significantly better over the years when observable classroom evidence is significantly improved. Skills developed are not being transferred to testing situations.

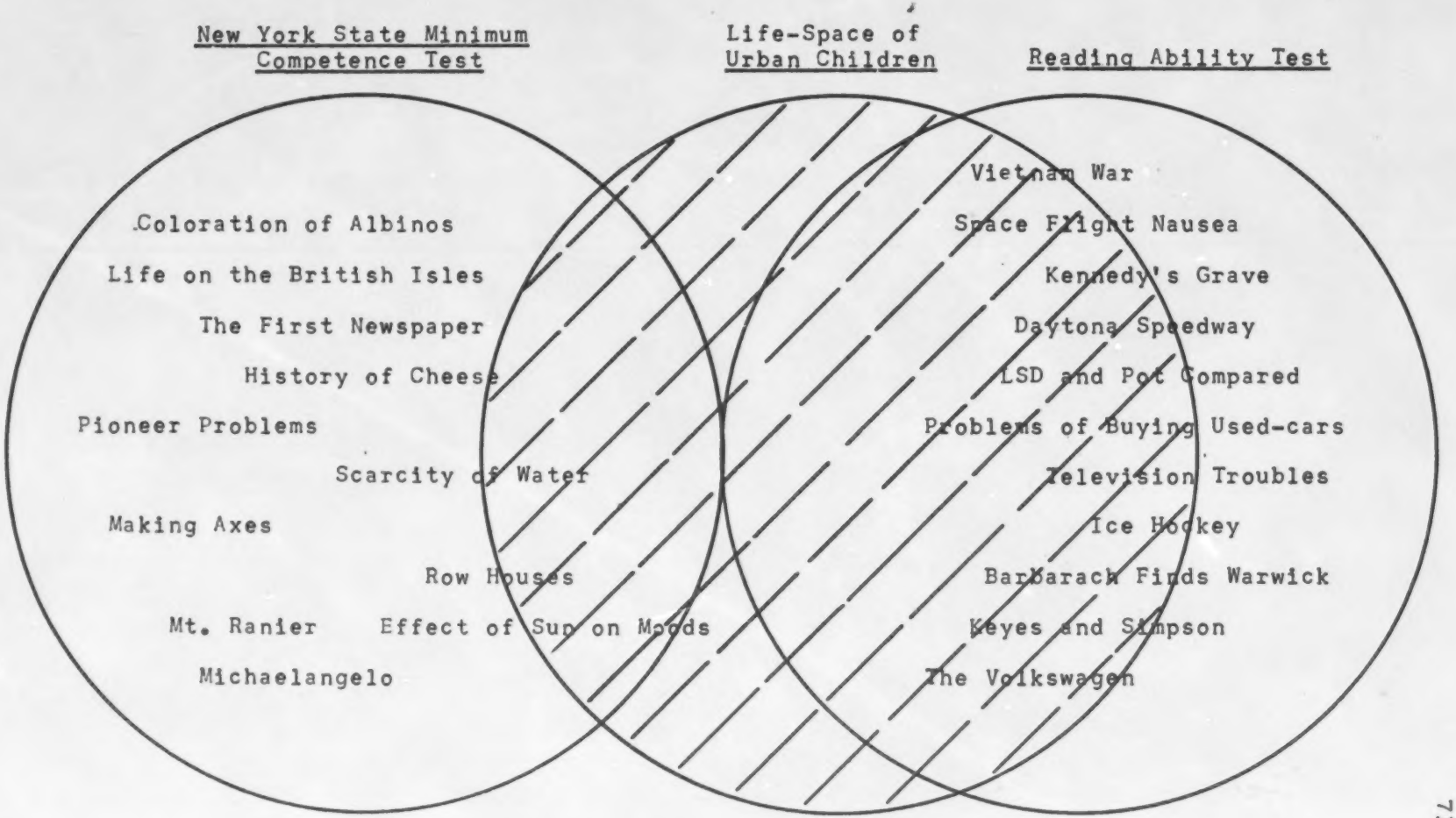
To investigate the nature of the lack of this transferrance, a critical examination was made of one test which fails to reflect encouraging classroom response. This was the Minimum Competence Test in Reading. It was conjectured that possibly the nature of the reading matter itself presented an obstacle to children who have minimal reading ability. Paragraphs were examined and found to have little connection with the realistic background in which these children live and learn. If, in addition to being asked to decode words, a child is simultaneously asked to cope with strange and unfamiliar concepts, he is in a sense, being tested for two distinct skills. Although this is perfectly justifiable in many testing instances, the question which arises here becomes: "Is it fair to test the minimal reading ability of children who have limited experiential background with a test which reaches beyond that background?"

In order to discover how subjects might respond to a test written within a realistic framework of experience, the Reading Ability Test was devloped. It was designed to be congruent with the Minimum Competence Test and in so far as possible, controlled all but the content factor. Maintaining similar format, the same degree of difficulty, close grammatical structure, a parallel flow of thought, and similar questioning patterns, the Reading Ability Test substituted relevant content in place of less relevant content. Figure 3 illustrates how this was accomplished. Children in urban areas are felt to have little first-hand knowledge and/or experience of many of the concepts on the Minimum Competence Test. They do not in everyday experiences generally encounter such things as: albinos, the British Isles, pioneers, axes, Mt. Ranier, or any of Michaelangelo's works. They do have some experience with cheese, water, row houses, and the sun, but the nature of paragraphs on the test does not insure that these concepts are readily conveyed.

Concepts on the Reading Ability Test are not proposed as being completely within the experiential range of the children in the experiment; however, these are topics with which an urban child is likely to have some association. The vicarious experiences of watching television and going to movies, together with such real experiences as listening to the conversation of older brothers, sisters, and peers is apt to have familiarized the subjects with most of the topics selected. As mentioned earlier, the topics have been sug-

Figure 3

A Comparison of the Reading Content of the New York State Minimum Competence Test in Reading and the Reading Ability Test as Each Relates to the Realistic Life-Space of Children in Urban Areas



gested for the study through the investigation of the real interests and values of ninth grade children in a summer Communications Course. This was an informal tapping of interests which was performed for a different purpose but supplied areas of authentic teenage interest.

The Reading Ability Test, available in Appendix B, is similar to the Minimum Competence Test in many ways. It was not, however, developed on as wide a scale as the Minimum Competence Test. It may have indirectly benefited from the State preliminary testing since items have been kept as similar as possible within the same framework. The test was first run with several individual children in grade nine in 1968. A revision followed and testing was extended to two seventh and two eighth grade classes in an urban area in the same year. It was repeated in two seventh grade classes in 1970, and extended to a whole ninth grade class in one school by this study.

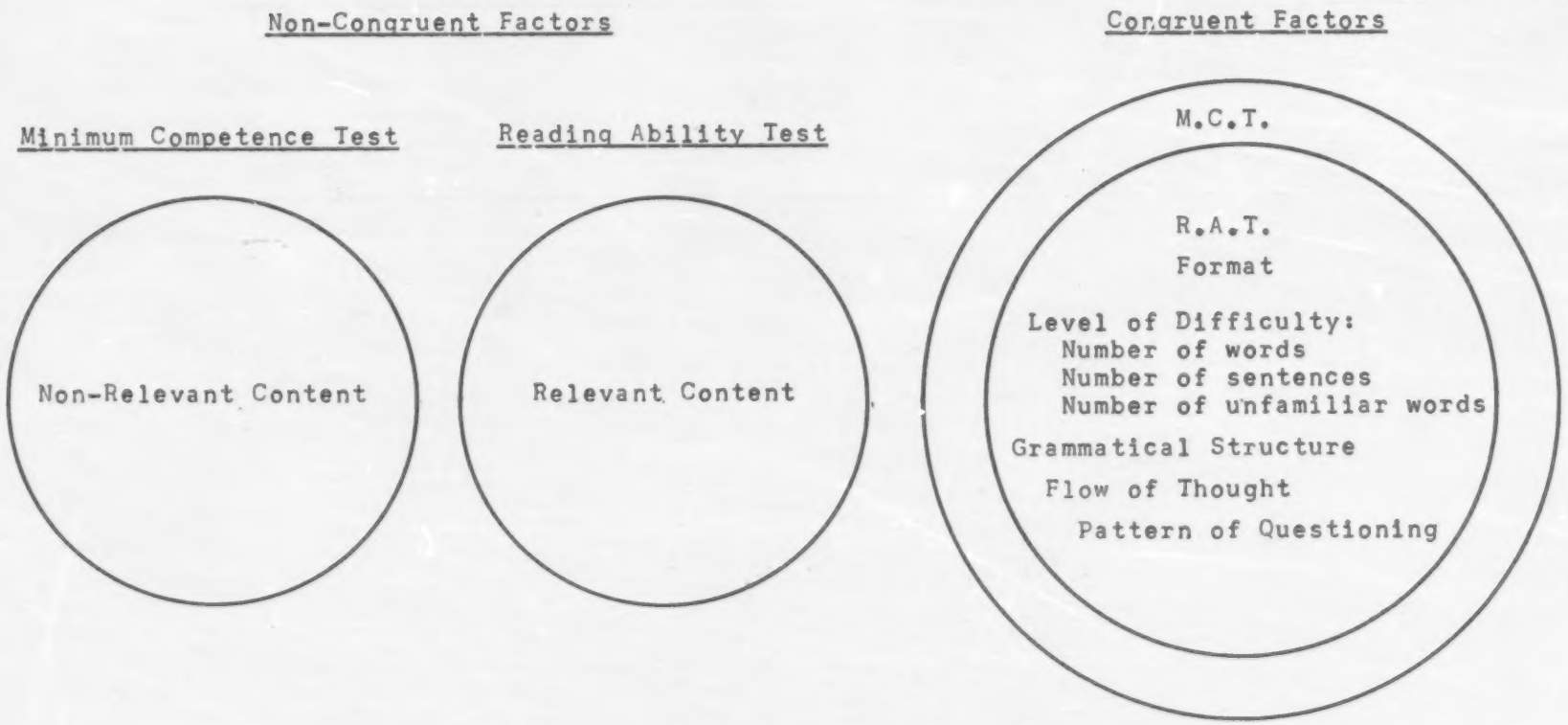
Congruency of Tests

Except for content, every facet of the New York State Minimum Competence Test in Reading has been kept similar in the Reading Ability Test. Figure 4 illustrates the facets maintained which are: format, level of difficulty, pattern of questioning, grammatical structure, and flow of thought. Each of these facets will be considered in this section in order to show how congruency was established.

Similar format was easier to achieve than the other elements of congruency. This amounted to the mechanical

Figure 4

Comparison of Congruent and Non-Congruent Factors in the New York State Minimum Competence Test (M.C.T.) in Reading and the Reading Ability Test (R.A.T.)



matching of printed forms. Permission was obtained from the Chief of the New York State Test Bureau of Pupil Testing to have the Minimum Competence Test in Reading changed from printed to typewritten form so that both tests would appear to be identical. The two tests were then run on an off-set press in order to obtain high quality copies. Matching answer sheets were printed for the Reading Ability Test so that these would parallel the New York State answer sheets. Copies of both tests and answer sheets may be reviewed in Appendices A and B. The same directions for both tests appear in the Appendix D. The answer key for both tests may be found in Appendix C.

A crucial point of similarity in the two tests was that of the level of difficulty of each test. If the Reading Ability Test was to be equally difficult, a statistical formula had to be applied first to the Minimum Competence Test to determine the readability level of each paragraph and then to the Reading Ability Test to establish that each paragraph was equally difficult. If children scored higher on the latter, it had to be the nature of the content which made the difference and not the fact that the congruent test was an easier test to read.

In order to establish matching difficulty, a formula of readability was selected and applied; the one chosen to

verify difficulty in this study was the Dale-Chall Formula of Readability.[1] This is a two factor formula which is familiar to professionals in the field of reading. It gives evidence of accurate statistical correlations; Dale and Chall[2] found, for example, that on fifty-five passages of health education reading material, the formula predictions correlated .92 with the judgments of reading experts and .90 with the reading scores of both children and adults exposed to the same material; material ranged from extremely easy to very difficult.

Three statistical facts enter into the application of the Dale-Chall formula; these are: number of words and sentences in a selection, together with the number of unfamiliar words as determined by comparison with a standard list of four thousand familiar words. In computing the difficulty level of each of the eleven paragraphs on both tests in this experiment, three people applied the formula to the paragraphs which guaranteed a triple check so as to insure accurate levels.

Table 1 illustrates the difficulty of each paragraph on both tests. The Reading Ability Test was revised repeatedly until this statistical consistency was reached. The range of both tests extends from fifth grade to fifteenth

---

[1]Edgar Dale and Jeanne S. Chall, "A Formula for Predicting Readability," Educational Research Bulletin, I (1962), 1-9, 11-28.

[2]Dale and Chall, p. 8.

Table 1

Dale-Chall Readability Rating of Paragraphs for the New York State Minimum Competence Test in Reading and the Reading Ability Test

Paragraph	Number of Sentences	Number of Words	Average Sentence Length	Number Unfamiliar Words	Dale-Chall Score	Grade Equivalent
A	9	100	11	21	7.50	9-10
B	4	90	23	5	5.57	5-6
C	5	134	27	20	8.13	11-12
D	7	117	17	7	5.63	5-6
E	9	122	14	8	5.59	5-6
F	7	157	22	28	9.15	13-15
G	6	105	18	29	9.11	13-15
H	5	68	14	6	5.28	5-6
I	11	190	17	26	8.59	11-12
J.	7	134	19	7	5.68	5-6
K	10	160	16	27	8.69	11-12

grade difficulty; the range is somewhat surprising for a test which is given to ninth grade children in order to determine their minimal reading ability. There appears to be a discrepancy between the State's classification of paragraphs and that established by this study. The State Test Manual observes that:

> In order to maintain the proper level of difficulty, the materials were checked against the Thorndike Word List; and words above seventh and eighth grade difficulty were replaced, except in those instances in which the item was testing meaning from context.[1]

Obviously, Thorndike and Dale-Chall differ significantly as to the classification of unfamiliar words. An examination of Appendix E illustrates the type of words considered unfamiliar according to the Dale-Chall standard. In addition, Dale-Chall consider average sentence length and number of words in a selection as integral to the difficulty of that selection. The New York State Test of Minimum Competence in Reading appears as a much more difficult test according to the Dale-Chall standard than according to the Thorndike List. The inference from this is that children are reading much better than currently believed.

It is essential to follow this line of reasoning one step further. Just what must a ninth grade child do in order to score minimum competence in reading according to the findings made through the Dale-Chall Formula? Table 2 illustrates his task. A child is said to have attained minimum

---

[1] N.Y.S. Education Department, p. 5.

Table 2

Reading Difficulty Range Chart for the New York State Minimum Competence Test in Reading and the Reading Ability Test

Paragraphs	Dale-Chall Grade Level	Number of Questions	
A	9-10	3	
B	5-6	3	
D	5-6	4	
E	5-6	3	21
G	5-6	3	
J	5-6	5	
C	11-12	3	
F	13-15	3	
G	13-15	3	19
I	11-12	6	
K	11-12	4	

Note: For a ninth grade child to score minimum competence in reading on either of the two tests, he must respond correctly, to at least 27 items.

competence in reading if he scores at least twenty-seven items correctly on either of the two tests in this study. According to the Difficulty Range Chart, for a child to achieve the competence score, he must be able to read five paragraphs (B,D,E,G,J) written below his grade level and answer all related questions correctly; in addition, he must read paragraph A, written at or slightly above his grade level and answer its related questions correctly. If he does this much accurately, he achieves a score of twenty-one which falls six points short of the minimum competence point which is twenty-seven. To arrive at that crucial point, he must then read five paragraphs which are written above his grade level (C,F,G,I,K) and answer at least six of the related questions correctly. Naturally, the law of chance may work in his favor allowing him to get six correct answers on difficult material. The question, however, is whether, in fact, a test of minimal reading ability should require a child to read and respond correctly to material written above his grade level and whether, in addition, the notion of minimum competence in reading should include a chance factor.

It is well to note the frustration of many urban children who must struggle to read up to grade level. A child, for example, who can cope with reading matter at his grade level but not above it may be fairly successful in dealing with paragraphs A and B on the tests. If paragraph C, written at an eleventh-twelfth grade level does not completely destroy his initiative, he may proceed to paragraphs

D and E and have his confidence restored since these are within his realistic ability range. By the time he reaches paragraphs F and G, the two hardest on the test, he is apt to be completely overwhelmed and may choose one of several options. Behaviors which follow have been observed by teachers in actual test situations where the New York State Minimum Competence Test in Reading has been given; urban children have been observed . . .

...giving up
...struggling on
...pretending they understand
...walking out
...marking answers indiscriminately
...cheating
...getting angry
...asking for help
...crying
...laughing
...sleeping

Unfortunately, in order to fulfill the purpose of this study, it was necessary to maintain the same unrealistic levels of difficulty on the Reading Ability Test. Children reading below grade level were faced with the same kind of frustration just described twice within a short period of time. The only point which somewhat justifies such a situation is that material on the Reading Ability Test is much more tolerable in nature despite the fact that some of it is much too difficult for ninth grade children. An item analysis of the paragraphs on both tests may lead to some interesting conclusions as to how these unrealistic reading levels affect achievement on specific items. It may prove that ability to cope with difficult reading material improves

with the relevancy of that material.

The third area of congruency in the two experimental tests is that of grammatical structure. The reader will be able to note over-all similarity of grammatical structure when he examines the two tests contained in Appendix A and B. However, a few examples are included at this point in order to illustrate where identical structure was possible and where tolerable deviations were allowed.

The first sentence of several matched paragraphs will be included here for examination; sentences typed in small letters represent sentences from the New York State Minimum Competence Test in Reading; sentences typed in capital letters represent the counterpart sentence on the Reading Ability Test.

Sample 1

Every kind of animal occasionally produces offspring that are white in color.

EVERY KIND OF SPACE CAPSULE OCCASIONALLY PRODUCES MOTIONS WHICH ARE FELT BY THE CREW.

In this sample, it was not possible to reproduce the word, animal, as one word in the congruent sentence; however, despite the fact that two words, space capsule, are used as a match, the two words contain an equal number of syllables; the match was therefore considered not ideal but acceptable.

There are samples where even this syllable agreement could not be achieved; this was due to the limitations put on flexibility by adherence to the Dale-Chall Readability Formula. It is doubtful that additional years of revision would estab-

lish a perfect relationship; however, the panel which ruled on the degree of congruency did not feel that the small number of remaining faulty matches interfered significantly with the total degree of congruency between the two tests. An example of a low degree of match between two sentences would be the following:

Sample 2

Even white lions, lobsters and giraffes can be seen if you visit the right place at the right time.

ROLLING, PITCHING, AND VIBRATING CAN BE FELT IF THE ASTRONAUTS PUSH THE WRONG BUTTON AT THE WRONG TIME.

Changing parts of speech as found in Sample 2 is hardly ideal on a congruent measure but sentences of this kind are in the minority, therefore, it was felt that they were not serious enough to deflect the purpose of the test as a whole.

The fourth element of congruency found in the two tests is flow of thought; this was achieved primarily by maintaining a similar grammatical structure. Initial sentences in each paragraph were relied on to set the matching train of thought in its counterpart, as in the following cases:

Sample 3

Riding among the first crude settlements of Indiana and Missouri, John Mason Peck preached the Gospel to anyone who would stop their work.

WORKING AROUND THE USED-CAR ESTABLISHMENTS OF NEW YORK CITY AND UPPER NEW YORK, JOE JOHNSON SOLD SECOND HAND CARS TO ANYONE WHO DESIRED TO TRADE.

Sample 4

There are places in the world where water is so scarce that it is almost worshipped.

THERE ARE PLACES IN THE WORLD WHERE TELEVISION SETS ARE SO SCARCE THEY ARE ALMOST WORSHIPPED.

Sample 5

For more years than anyone can remember, the American ideal of a proper home has been a house with a yard around it.

FOR MORE YEARS THAN ANYONE CAN REMEMBER, THE AMERICAN IDEAL OF YEAR-ROUND SPORTS HAS BEEN BASEBALL IN THE SUMMER AND FOOTBALL IN THE WINTER.

The final point at which the two tests in this experiment are considered congruent is the pattern of questioning. As with matching paragraphs, the question construct was adhered to as closely as possible for each item. Whenever possible, exact types of questions were asked, for example.

Sample 6

Albinos must guard against

1	Africans	3	cold weather
2	lions	4	the summer sun

SPACE CREWS MUST GUARD AGAINST

1	LONG FLIGHTS	3	COLD WEATHER
2	FEVER	4	CAPSULE MOVEMENTS

In the item construction, as with paragraph construction, matches are as accurate as possible and sufficiently close to make a meaningful comparison. It was considered advisable in at least one or two instances to use an identical vocabulary item within the two distinct types of content, for instance:

Sample 7

(Item 22)

As used in this story, the word "managed" means

1 liked 3 studied
2 gave orders 4 was able

AS USED IN THIS STORY, THE WORD "MANAGE" MEANS

1 WOULD LIKE 3 STUDY
2 GIVE ORDERS 4 ARE ABLE

So as not to expose the children to two distinct sets of context clues which on the second test taken might unlock the meaning whereas on the first test it failed to do so, most context questions substituted new, rather than the same, vocabulary words as in the following case:

Sample 8

(Item 8)

The phrase "coping with" is closest in meaning to

1 discovering 3 solving
2 handling 4 guiding

THE PHRASE "BARGAINING WITH" IS CLOSEST IN MEANING TO

1 DISCOVERING 3 STOPPING
2 TRADING 4 GUIDING

This concludes the comparison of elements of congruency which were built into the Reading Ability Test so that according to format, level of difficulty, grammatical structure, flow of thought, and pattern of questioning it would be externally similar to the New York State Minimum Competence Test in Reading.

Nine professionals were asked to rate the congruency of the Reading Ability Test in relation to the New York State

Minimum Competence Test in Reading; represented in this group were: two reading specialists, two teachers of reading, one social case worker, one social psychologist, one social studies teacher, and two teachers of English. The Rating Sheet used by panel members is included in Appendix I.

In addition, the Chief of the New York State Bureau of Pupil Evaluation reviewed the test and suggested specific changes to increase its accuracy.

Most paragraphs were rated within the acceptable to acceptable-with-modification range. Some recommendations for change had to be ignored since they would have required such major alterations that parallel grammatical structure, for example, would have been destroyed. Keeping the Reading Ability Test congruent with its counterpart necessitated maintaining some of the weaknesses in the Minimum Competence Test; however, more basic changes would have defeated the purpose of the study. As mentioned earlier, the Reading Ability Test was not proposed as a necessarily improved version of its counterpart, merely, as a vehicle for testing a theory concerning content.

Content of Tests

The substitution of relevant content in place of content which is alien to the immediate background of urban children formed the basis of this study. It was maintained that testing the minimal reading ability of urban children on unfamiliar material is testing them unfairly. There is, how-

ever, an even more crucial question involved in the nature of reading material presented to urban children. Leacock, who has done extensive research among children in metropolitan areas, claims that

> . . . the content of school readers can utterly devalue the experiences of Negro and working-class children by erasing them from the world; it poses the problem that in order to accept formal educational goals that involve mastering such material, they must devalue themselves.[1]

This study extends the notion of devaluation to the testing process. Both Raths[2] and Ozman[3] are also convinced of the tremendous need for clarification of values transmitted in education; Raths' concern is more with developing the processes of value clarification among children; whereas, Leacock[4] and Ozman[5] are more concerned with the need for educators to investigate what values are being transmitted through all types of reading materials which are utilized by children at every level of education.

It would seem that tests, in particular, deserve special scrutiny since the test situation is perceived by children as important. They perceive that adults attach

---

[1]Eleanor Burke Leacock, Teaching and Learning in City Schools: A Comparative Study (New York: Basic Books, Inc., 1969), p. 80.

[2]Louis E. Raths, Values and Teaching (Columbus, Ohio: Charles E. Merrill Publishing Company, 1966), pp. 6-7.

[3]Howard A. Ozman, Jr., "Value Implications in Children's Reading Material," The Reading Teacher, XXII (December, 1968), 249.

[4]Leacock, p. vii. [5]Ozman, p. 249.

great importance to a test situation; therefore, what appears in a test must be valuable. It is not maintained that urban children necessarily perceive the test situation in this manner; however, there is the possibility that they may. The test administration process itself conveys the notion that what is being tested in such a formal manner must in itself be important. Tests of reading, therefore, deserve careful examination particularly in regard to the danger of devaluation of a child's experiences.

The Reading Ability Test eliminates, or at least, reduces the problem of devaluation since it was based immediately on the values, interest, and experiences of urban children. In order to arrive at suitable content, one major source of information was tapped, namely, the children themselves.

Pupils' interests were sampled directly through a six weeks summer program in Communications held at the Albany High School Annex in 1968. This was a Title I project designed to improve the reading skills of ninth grade children through a language-experience approach. Twenty teachers in the program were given orientation in the language-experience method; in addition, they were given copies of Raths' book, Values and Teaching, which was designed to direct teachers in making children aware of what they value and how they came to value it. Appendix H illustrates some of the tools used by

---

[1]Raths, pp. 3-258.

the teachers who helped children in the value clarification process. Information on student interests collected through the Communications Course became the basis for content of the Reading Ability Test. The most prominent areas of general interest identified at the end of summer 1968 were:

Sports	Being Accepted
School Problems	Television
Cars	Vietnam
Family Relations	Sex
Marks	Movies
The Future	Friends
Popularity	Drugs

Content selected for the Reading Ability Test as contrasted with that appearing on the Minimum Competence Test appears on page 71, Figure 3. Content selected was considered to be within a realistic life-space of the child since it was identified by children themselves. Havighurst[1] defines life-space as the physical space, the objects contained within that space, and the people who inhabit that space in which the child exists. An important factor within that space today is television; specifically, programs which are identified as important to urban children. Therefore, reading content which may appear to be outside the actual life-space of a block-oriented child may include such things as space travel, sports, and national figures since these are familiar to him through the medium of television.

---

[1]Robert J. Havighurst, and Bernice L. Neugarten, Society and Education (Boston: Allyn and Bacon, Inc., 1962), p. 206.

SUBJECT POPULATION

Since the New York State Minimum Competence Test in Reading is mandated for all ninth grade children, choice of grade level for the experiment was restricted. Ninth grade children in Livingston Junior High School were selected for the study because Livingston is located in a typical, depressed urban area and houses all but a few ninth graders in the North End of Albany. St. Joseph School enrolls the remainder of ninth graders in the North End but including them in this study would have introduced another variable; namely, the non-public school element. Since no significance is attached by this study to how content affects religious-oriented as opposed to non-religious-oriented education, St. Joseph School was eliminated. The other public school in Albany which could have been included was Hackett Junior High School; the school, however, is less typical than Livingston in that it houses a large advanced placement program. Since the immediate concern of this study was with children who are apt to score at the lower end of a reading scale, Hackett was also eliminated. Children enrolled in private schools and institutions were also considered inappropriate subjects for the study.

There are 187 children enrolled in the ninth grade at Livingston but on the test day only 151 were in attendance; this group constituted the total number included in this experiment. Of this group, 75 were boys and 76 girls; 69 were

black and 82 white. These divisions together with eight class units formed the basic sub-groups in the study.

TEST PROCEDURE

Groups were randomized so as to facilitate data analysis procedures. The tests were administered on the morning of October 8, 1970. One half of the total group took the New York State Minimum Competence in Reading Test first and the second half took the Reading Ability Test first. At the end of the forty minute test period a brief respite was provided and the testing process reversed.

Prior to the test date, the principals and the two guidance counselors in charge of the testing were instructed in the conditions of the experiment. They, in turn, met with teachers from the eight participating classes who shared in the administration of the tests. Teachers assisted with the distribution of test booklets, made sure that pupils did not open test booklets or turn pages until instructed to do so, and collected materials in the proper order at the close of each test.

Scoring was done at the Board of Education where it was possible to double check results; hand stencils were used in this process. Results of the New York State Minimum Competence Test in Reading were returned to the school and filed with the State Education Department.

Chapter 4

PRESENTATION AND ANALYSIS OF THE DATA

This chapter is concerned with the presentation and analysis of data which will be used in determination of the significance of the experiment. Such data will include: test of the null hypothesis, correlation and alienation coefficients, four-fold distribution, mean score variations of sub-groups, reliability coefficients, and item analysis.

TEST OF THE NULL HYPOTHESIS

Mean Differential

In order to determine the probability that observed mean score differences did not occur merely by chance, the null hypothesis was tested at the .05 level of confidence. Comparison of the mean differences in performance between the Minimum Competence Test and the Reading Ability Test produced a z score of 2.3 which, although slightly below the .01 level of confidence is well within the .05 level. On this basis, the null hypothesis was rejected and the differences considered significant.

Coefficients of Correlation and Alienation

The Pearson Produce-Moment Coefficient of Correlation (r) was computed. The r = .685, with the coefficient of

determination r^2 = .469. Application of the test t to r^2 indicated that the relationship of .469 was significantly different from zero at the .01 level and could not have happened by chance.

The coefficient of non-determination k^2 which is the compliment of r^2 was determined to be .531; this is simply computed by the formula $k^2 + r^2 = 1.00$; k^2 indicates the degree to which results of the Minimum Competence Test and the Reading Ability Test are unrelated. K or $\sqrt{1 - r^2}$ provides a coefficient of alienation of .728 supporting a significant lack of relationship.

Application of the t test to the k coefficient produced a t score of 12.9 which is well within the .01 level of confidence limits. Using the data available with an N = 151 and 2 degrees of freedom, a .01 level of confidence can be achieved with a t = .288.

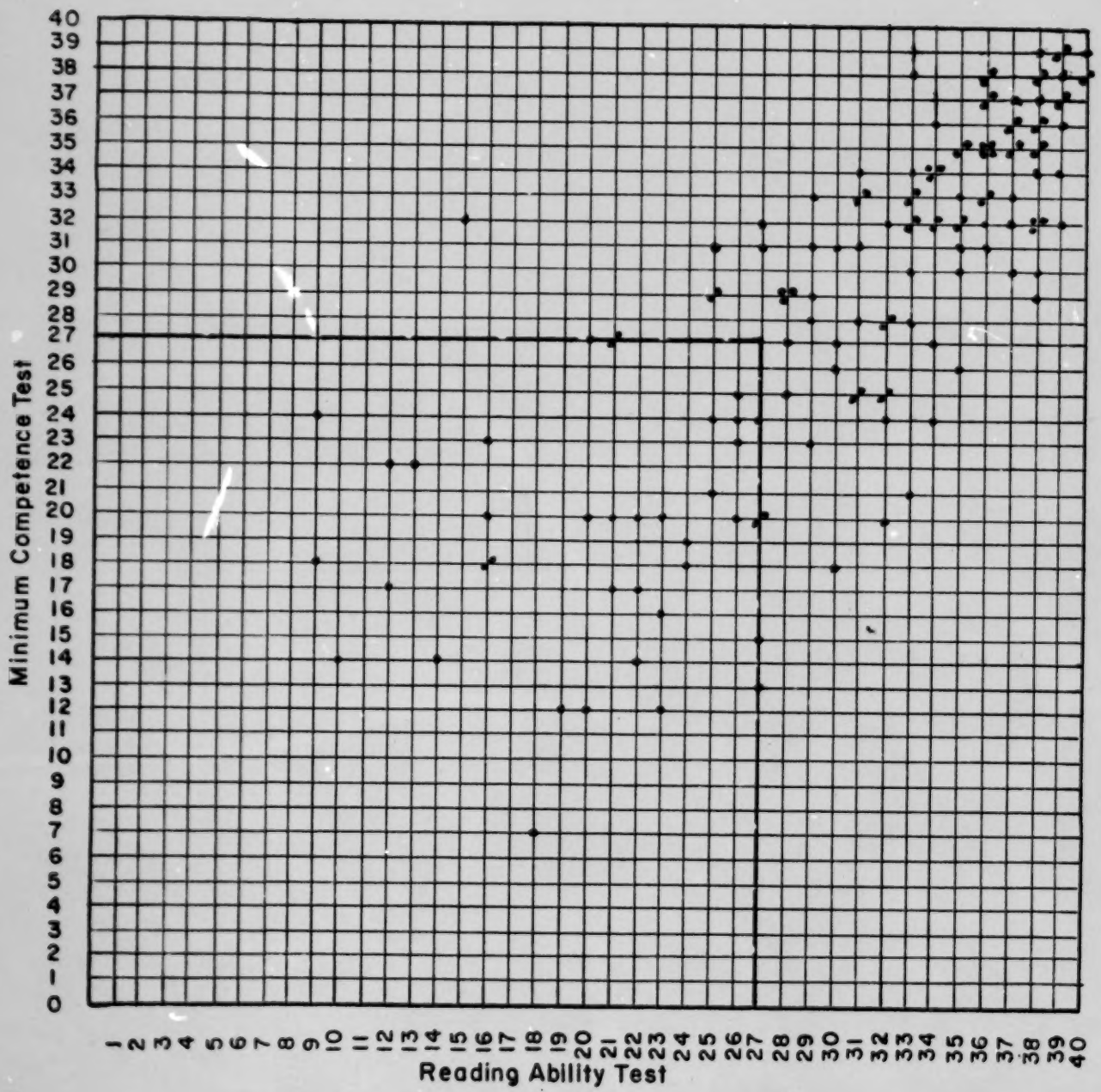
The Scattergram, Figure 5, illustrates the strong lack of relationship found between results of the Minimum Competence Test and the Reading Ability Test. Guilford[1] points out that "Sometimes we wish to stress the point of independence between two things rather than their closeness of agreement. In such instances, we present k as well as r." Such was the purpose of this analysis: to determine the extent of dissimilarity which two types of content created

---

[1]J. P. Guilford, Fundamental Statistics in Psychology and Education (New York: McGraw-Hill Book Company, Inc., 1956), p. 376.

Figure 5

SCATTERGRAM



within the same group of subjects. The Scattergram shows that the .685 correlation was undoubtedly strongly affected by the performance of the better readers. Apparently, the nature of content has little effect on children who appear in the upper right hand quadrant of the Scattergram. Superior readers function equally well on both relevant and non-relevant reading matter.

The k coefficient of .728, on the other hand, illustrates the dissimilar reaction of less capable readers. Apparently, the nature of content affects them considerably as demonstrated by the inverted cone shape of the Scattergram. Scores are closely clustered between the 30-40 range but widely dispersed between the 10-30 range. The spread of scores found below the twenty-seven minimum competence point on both tests indicates that two distinct types of reading content produced significantly different results.

Since the concern of this study was with children who scored below minimum competence on either or both tests, more specific observations of this group are appropriate at this point. A statistical analysis of them independently of the total group might result in biased data; however, it is possible to investigate individuals in this quadrant in terms of their history of reading achievement and their individual performance record in this experiment.

Children who scored below minimum competence in reading on both tests in the experiment were identified on Figure 5, the Scattergram, as those falling within the dotted line area enclosure.

This group is composed of both borderline cases who are close to, but not within the minimum competence range; it also represents severely disabled readers who, in spite of years of instruction in reading, continue to have difficulty with the decoding process.

In order to determine more specifically how children in this low scoring section performed, an examination was made of their individual responses to specific items. The items to which they responded best on both tests (60%-90% correct) were identified. In addition, items on which they scored lowest (12%-59% correct) were also identified; these items are as follows:

Table 3

Items Eliciting Best and Poorest Responses on the Minimum Competence Test and the Reading Ability Test

List 1	List 2
Best Responses on Both Tests	Poorest Responses on Both Tests
No. 1	No. 17
3	19
4	26
9	30
10	35
15	37
16	38
22	
23	
33	

There is a marked difference in the reading skills needed to deal with these two sets of questions. The items on List 1

are of a literal-fact-recognition type. Relatively simple reading skills are required for these items. They deal with straight-forward facts; such as, what a word means from context; what situation or condition exists; or how two things obviously compare.

List 2, on the other hand, requires more sophisticated skills of reading since the items deal with central thoughts, chief advantages, mathematical facts, metaphorical meanings of words, and the literary classification of a selection. As might be expected, less capable readers are unable to cope with the more sophisticated reading tasks which such kinds of items demand.

The important observation, however, is that the children in question, do actually possess true minimal ability in reading in the sense that they can deal successfully with the least sophisticated of reading skills; namely, simple identification of word meaning and factual recall. They are apparently non-competent beyond this level but have evidenced a real ability to read and comprehend certain types of items. This ability cannot be attributed to chance because the pattern of response is so apparent. That pattern is illustrated by the fact that among children in this low scoring section, 60 per cent to 90 per cent of them answered correctly one set of matching questions on both tests. Only 12 per cent to 59 per cent of the same group succeeded in answering more complex matched questions on both tests. They may be said to be competent in responding to basic, concrete, factual reading tasks, but non-

competent in relation to reading which requires more sophisticated skills. This is quite different from labelling them as generally non-competent in reading which, unfortunately, had been done on school records.

The intelligence range of children in this group has been investigated and, as suspected, is low, encompassing children whose scores run from 62 to 107. However, since the intelligence tests are principally verbal in nature, the same problem of alien content deflecting true achievement may be reflected in these intelligence scores. Intelligence scores may prove higher on test matter which is based on a realistic experiential background.

One final observation must be made of children in the "failing" group and that is the fact that four of them answered 96 per cent to 100 per cent of the items attempted correctly but the allotted time of forty minutes proved insufficient for them to complete more than twenty to twenty-four items. These very accurate but painfully slow readers are considered non-competent despite the fact that they made no mistake in answers attempted.

OTHER SUPPORTIVE DATA

Four-Fold Distribution of Compared Scores

Table 4 illustrates the percentage of pass-fail combinations which resulted from both the Minimum Competence Test and the Reading Ability Test. Results indicate that 60 per cent of the 151 subjects tested passed both tests; 21 per cent

Table 4

Four-Fold Distribution of Compared Scores
Minimum Competence Test in Reading -
Reading Ability Test

Re: Pass-Fail Relationships

Minimum Competence Test	Reading Ability Test: Pass	Reading Ability Test: Fail	Total
Pass	N 90 60%	N 7 5%	N 97 65%
Fail	N 21 15%	N 33 21%	N 54 35%
Total	N 111 74%	N 40 26%	N 151 100%

failed both tests; 14 per cent passed the Reading Ability Test but failed the Minimum Competence Test and 5 per cent passed the Minimum Competence Test but failed the Reading Ability Test.

The most important conclusion which these relationships indicate is that twenty-one children, or 14 per cent of the total group tested, are children who, at this writing, have been classified as non-competent readers. Their failure on the Minimum Competence Test is already recorded on local and State report forms and they are being tallied into profiles as non-competent. However, on an equally difficult test which was written within a realistic framework of their experience, they have demonstrated competence in reading. They were apparently unable to cope with unfamiliar concepts and, as a consequence, have been labelled as failures in reading.

An important question raised by the Four-Fold Distribution is that of the seven children who passed the Minimum Competence Test but failed its more relevant counterpart. Five of these took the Minimum Competence Test first in the order of testing and it is feasible that fatigue and/or resentment at having to take a second test immediately may have produced this unanticipated result.

Sex and Race, Mean Score Comparisons

When it was determined that the total group scored significantly higher on relevant reading material, it was decided to examine several sub-groups. The purpose of such

examination was to discover if relevant content affected boys more than girls and blacks more than whites. In addition, the eight participating classes at Livingston Junior High School were investigated as to the similarity of their response.

Table 5 considers the mean score differentials between boys and girls and blacks and whites on both tests of reading. As anticipated, all groups showed an increased mean score on the Reading Ability Test as compared with the Minimum Competence Test. Black girls and white boys responded most positively to more relevant content as evidenced by a +2.4 and +1.9 increase respectively. Black boys and white girls responded better on relevant material also but not to the same degree. Although overall gains were made by each of these sub-groups, no one group made highly significant gains over any other group. The response of the group as a whole to relevant material is much more significant than the response of any of these sub-groups when compared with each other.

Comparison of Individual Classes

The overall significant change pattern which statistical analysis produced may be seen graphically in this section. The distribution charts which follow illustrate the trend of improvement which relevant content produced. The dotted line on each chart is the minimum competence point of twenty-seven. Classes have been arranged and arbitrarily designated by letters. Class A made the greatest gain and class

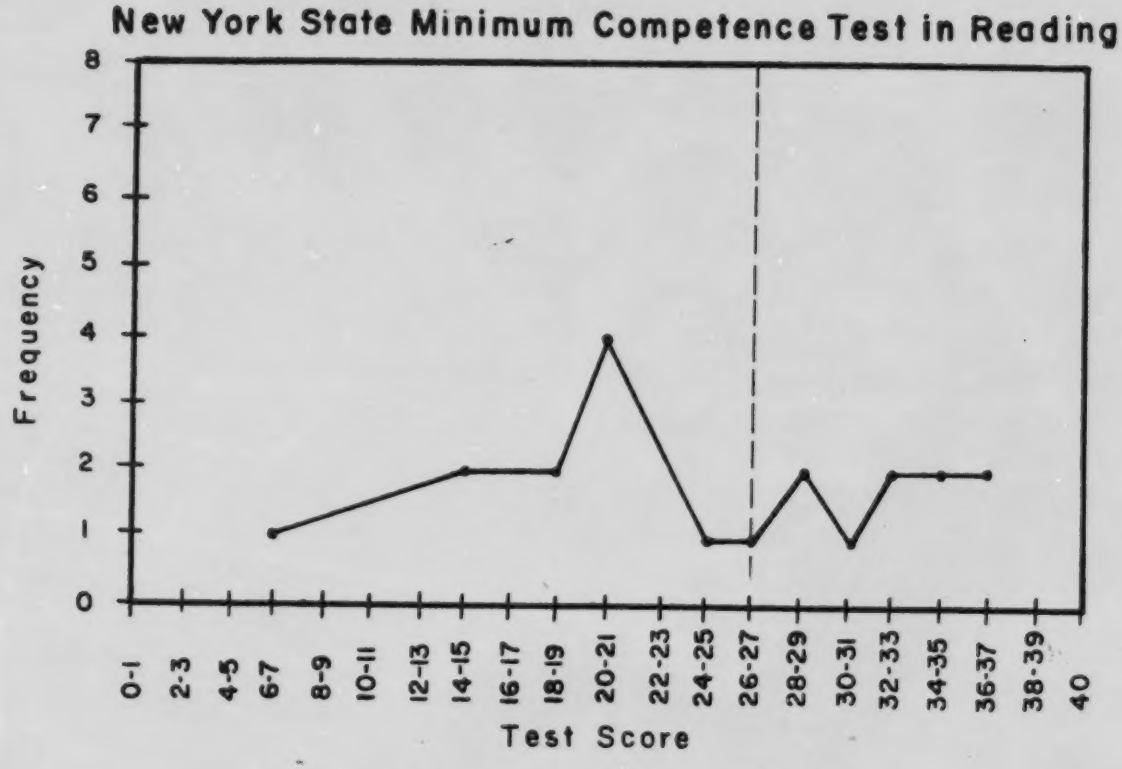
Table 5

Mean Score Comparisons of Sex and Race Factors in Relation to Relevant Reading Material

Sex	Race	Minimum Competence Test		Reading Ability Test		Differential
		Mean	SD	Mean	SD	
BOYS	Black N 26	24.4	8.2	25.7	7.8	+1.3
	White N 46	29.4	7.7	31.5	8.1	+1.9
	Combined N 72	27.8	7.8	29.3	8.4	+1.5
GIRLS	Black N 43	25.6	7.2	28.0	7.7	+2.4
	White N 36	30.9	6.3	32.3	6.5	+1.4
	Combined N 79	28.1	7.3	30.2	7.0	+2.1

Figure 6

Frequency Polygons-Class A



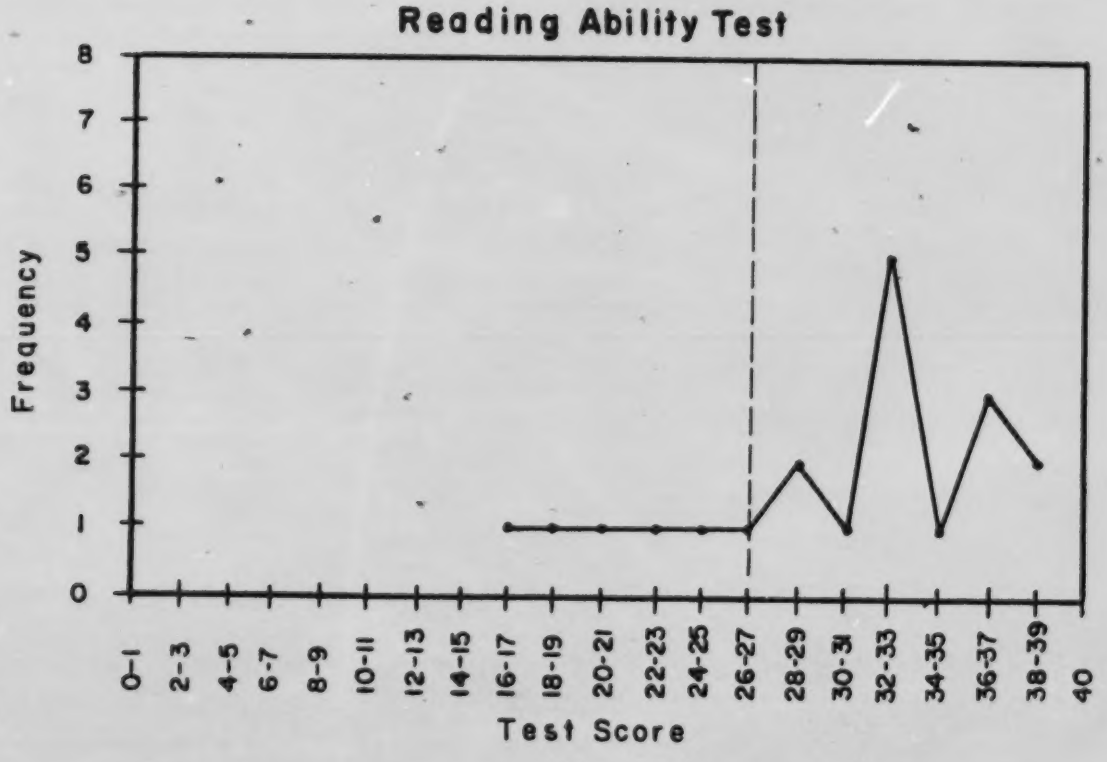
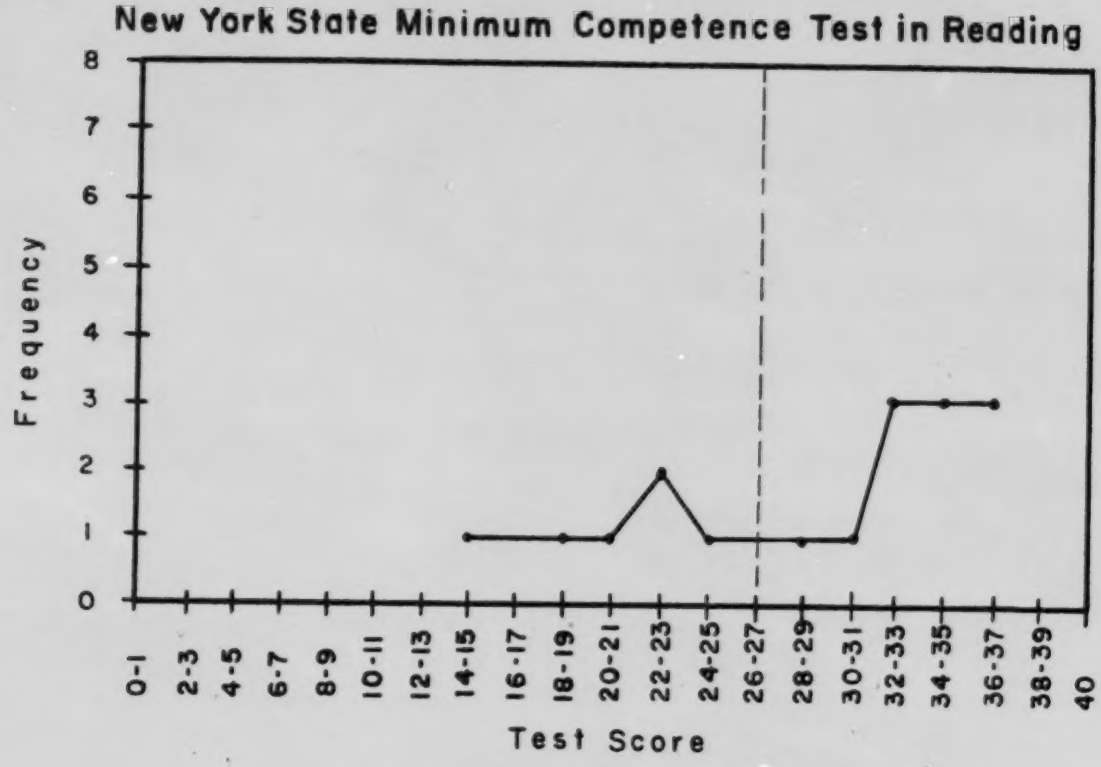


Figure 7

Frequency Polygons-Class B



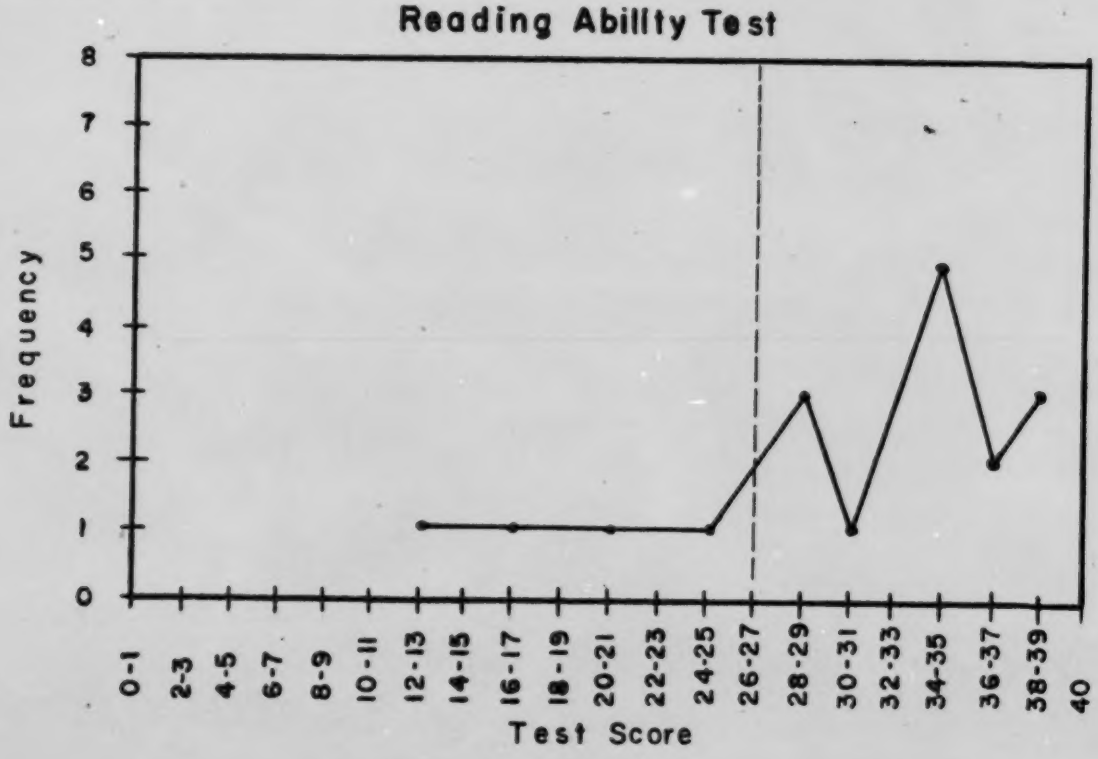
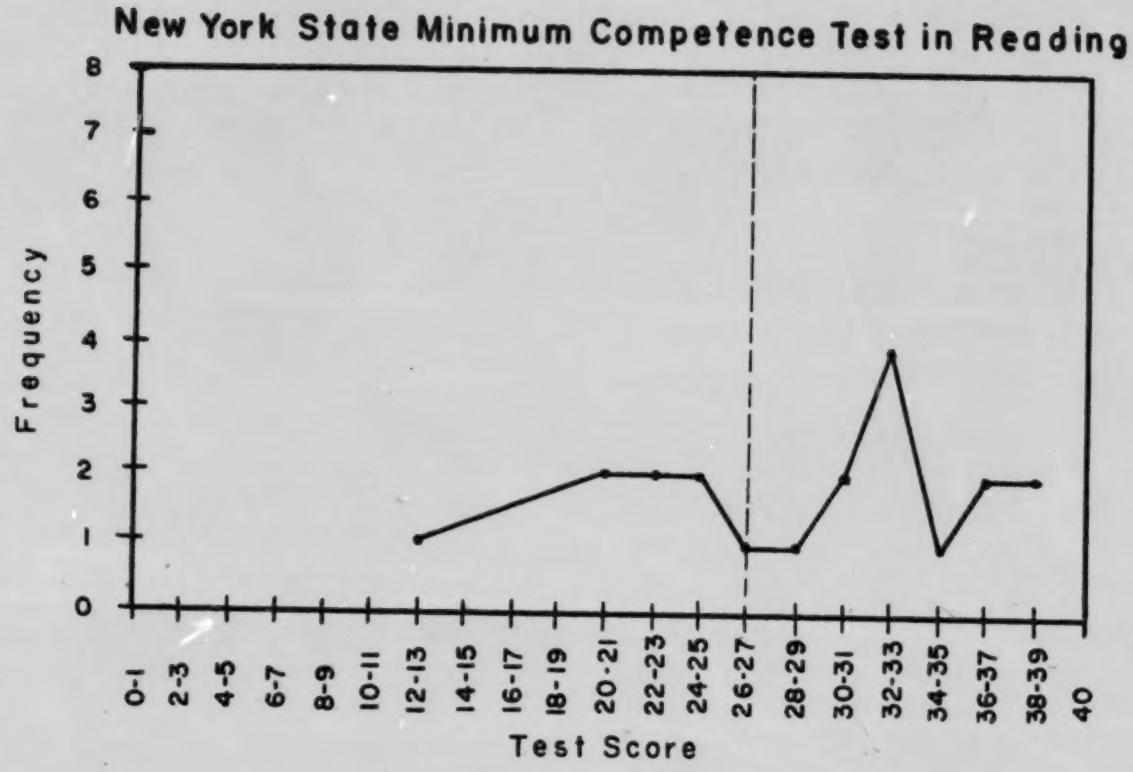


Figure 8

Frequency Polygons - Class C



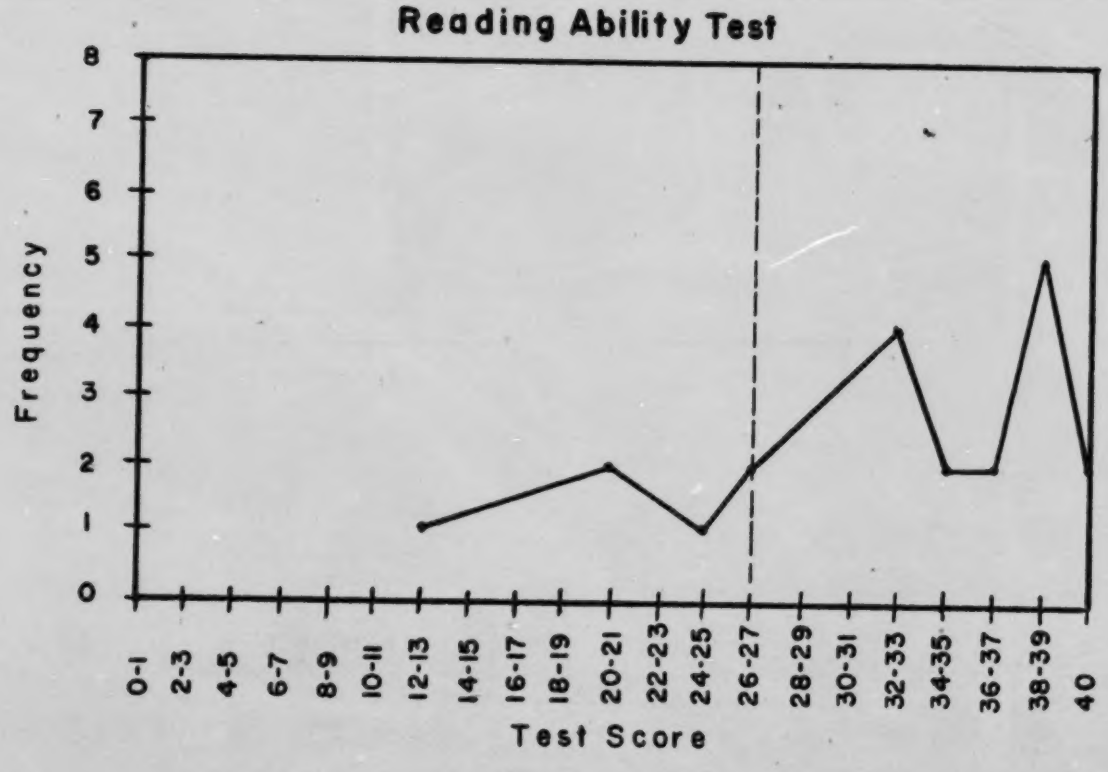
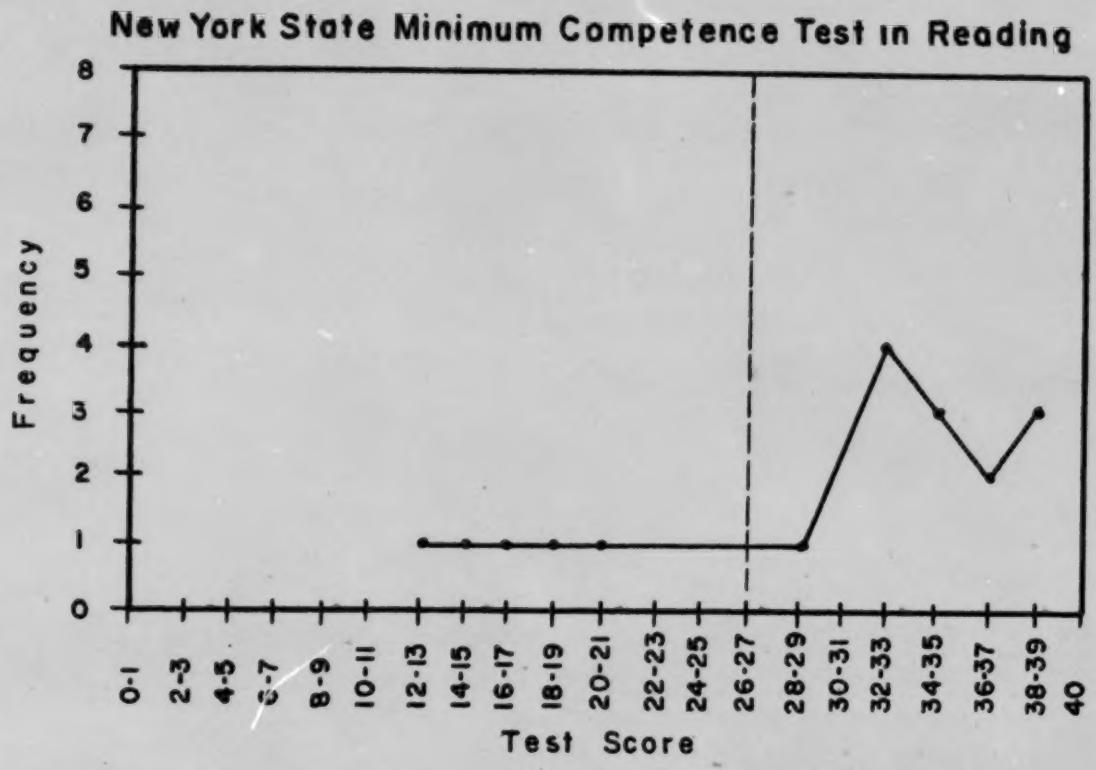


Figure 9

Frequency Polygons-Class D



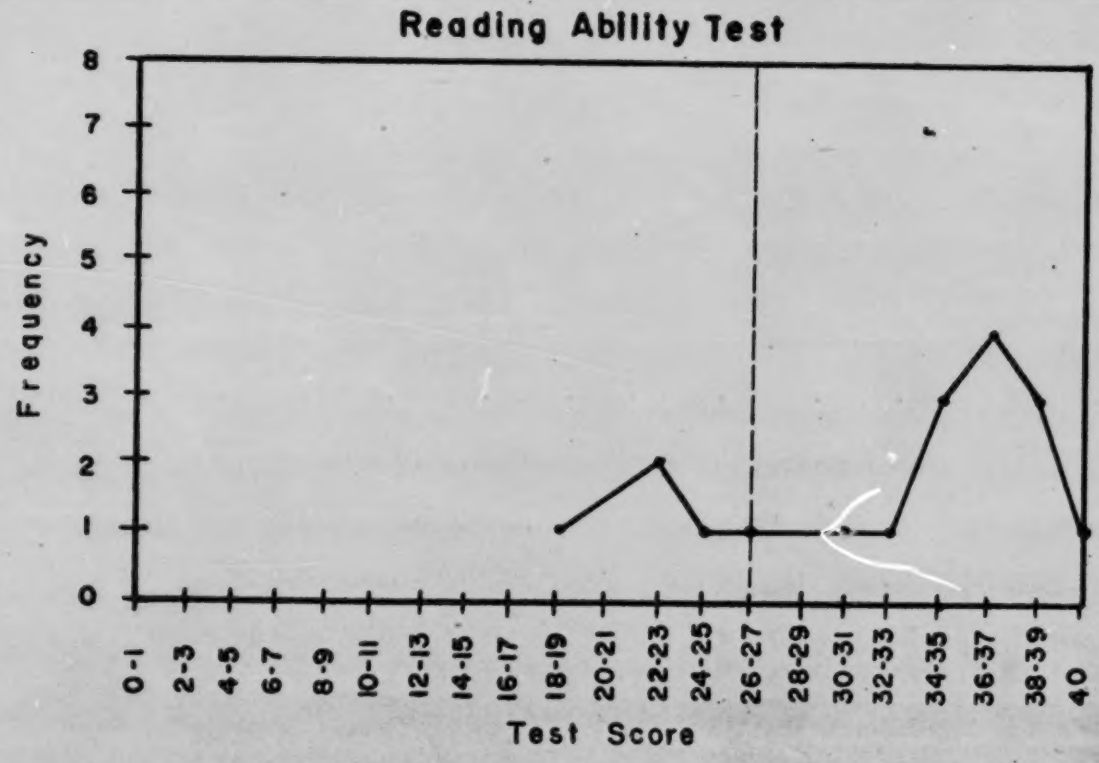
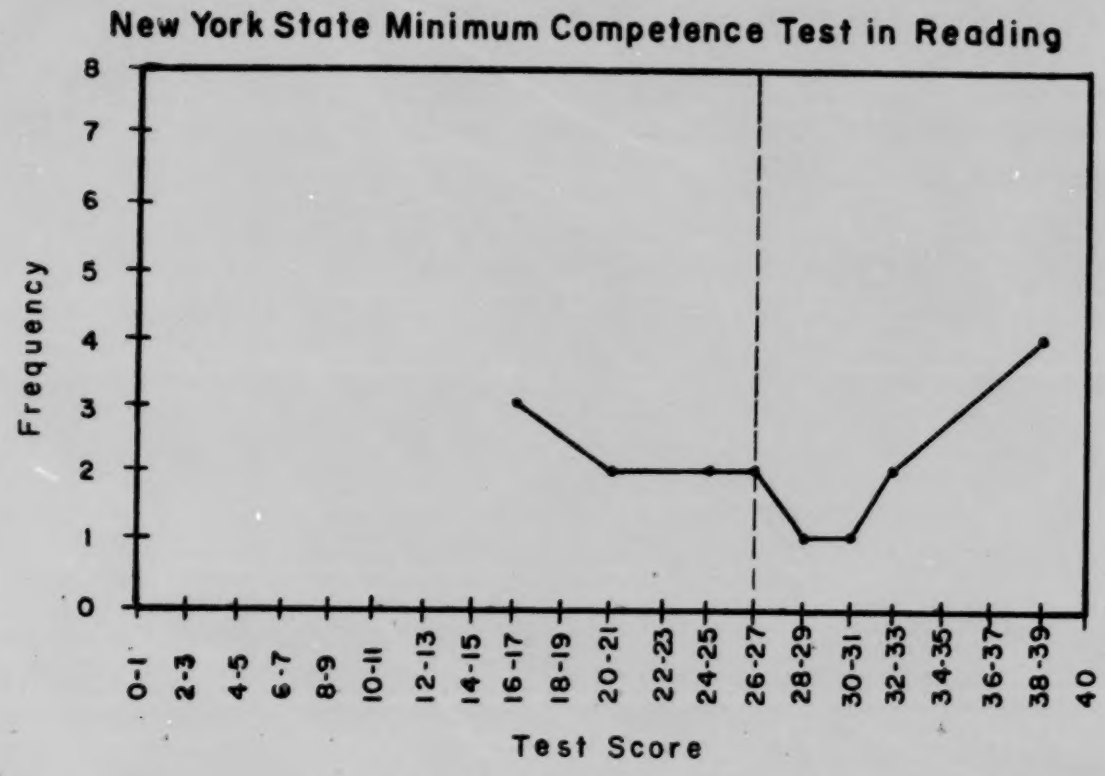


Figure 10

Frequency Polygons-Class E



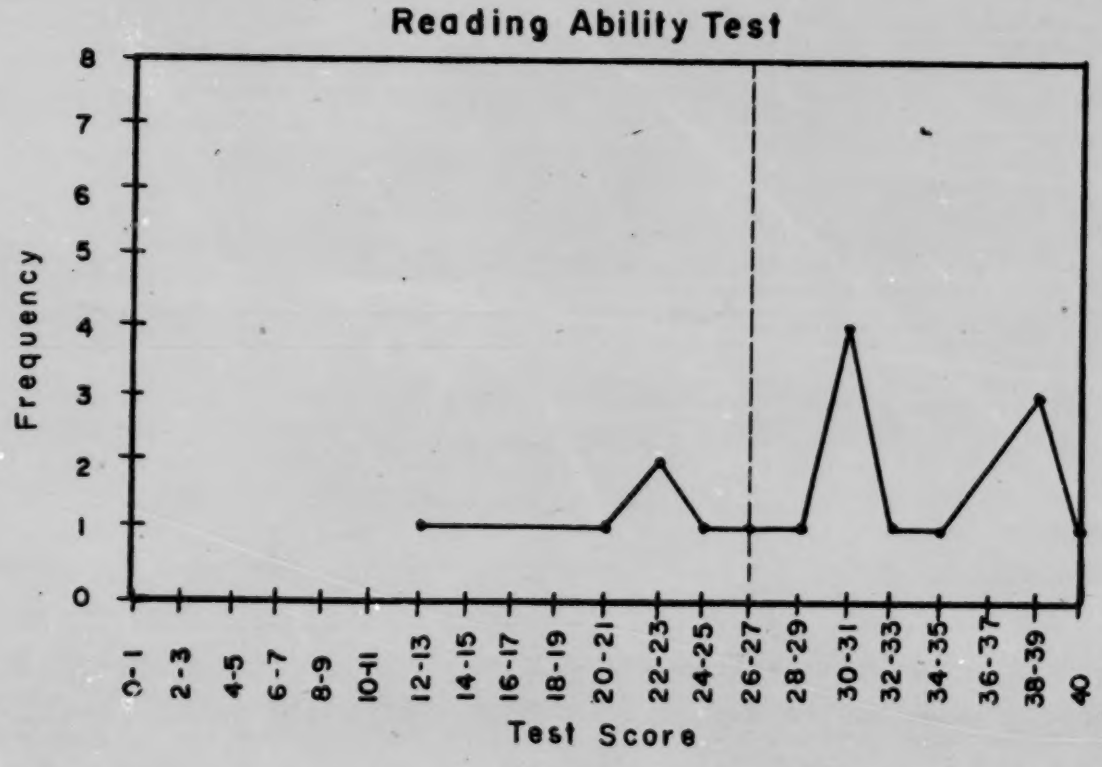
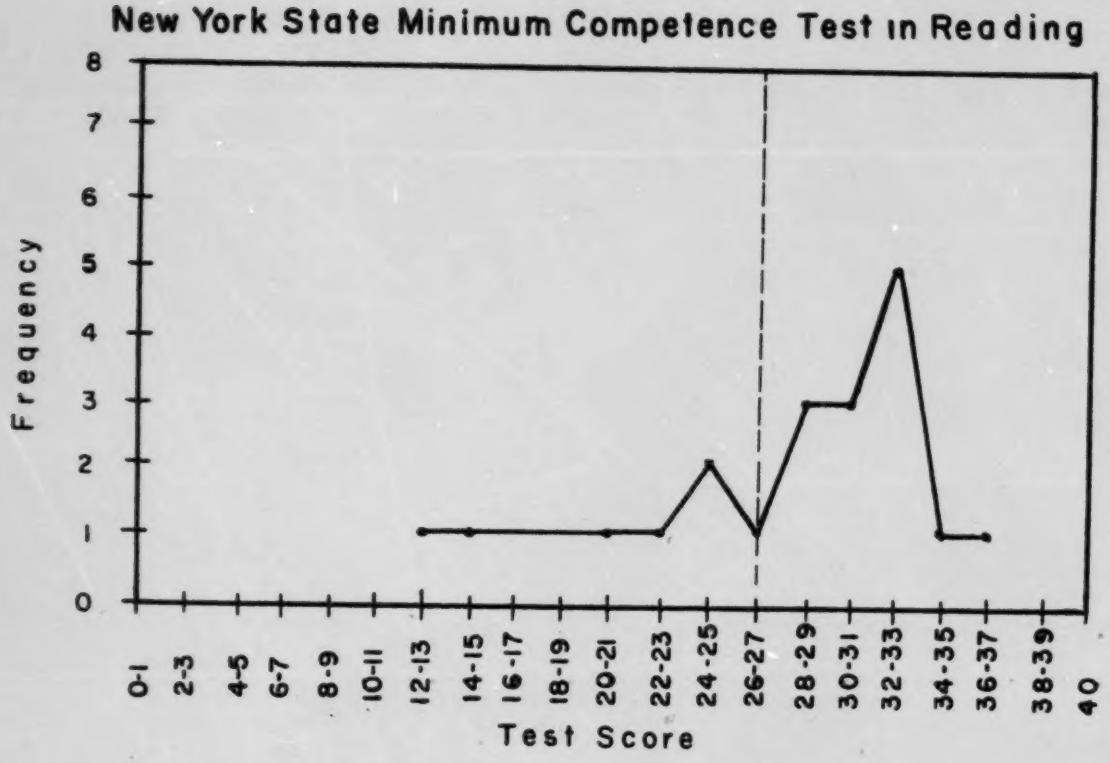


Figure II

Frequency Polygons-Class F



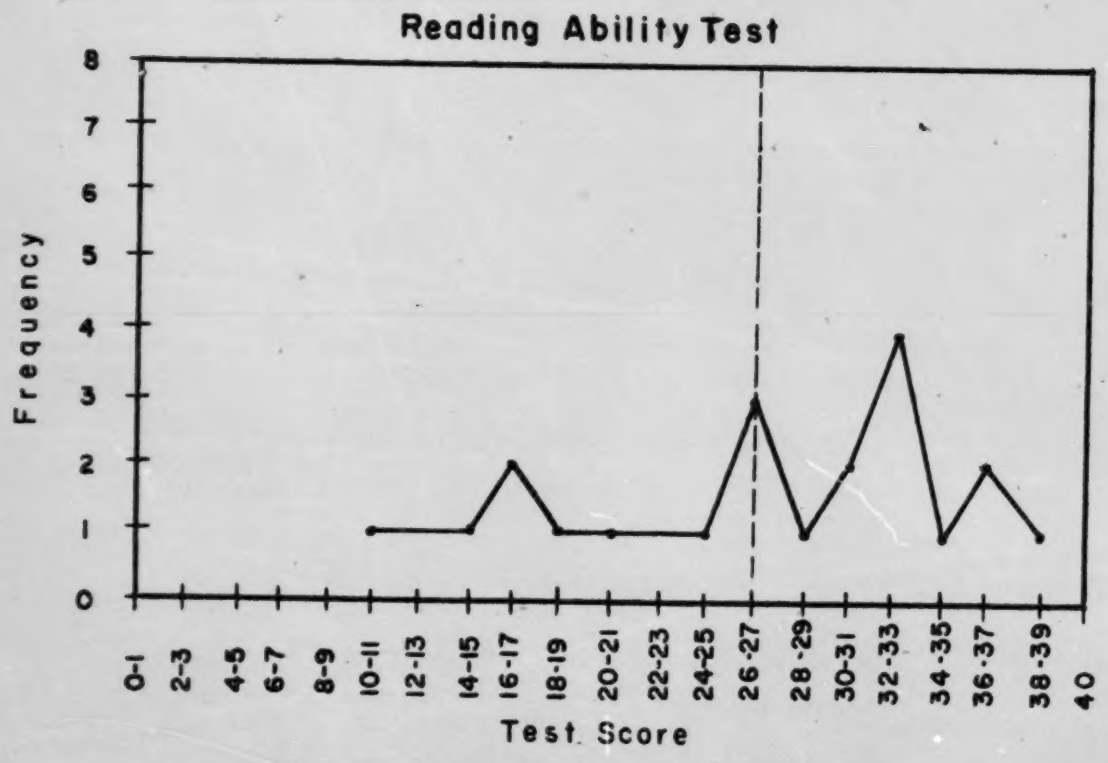
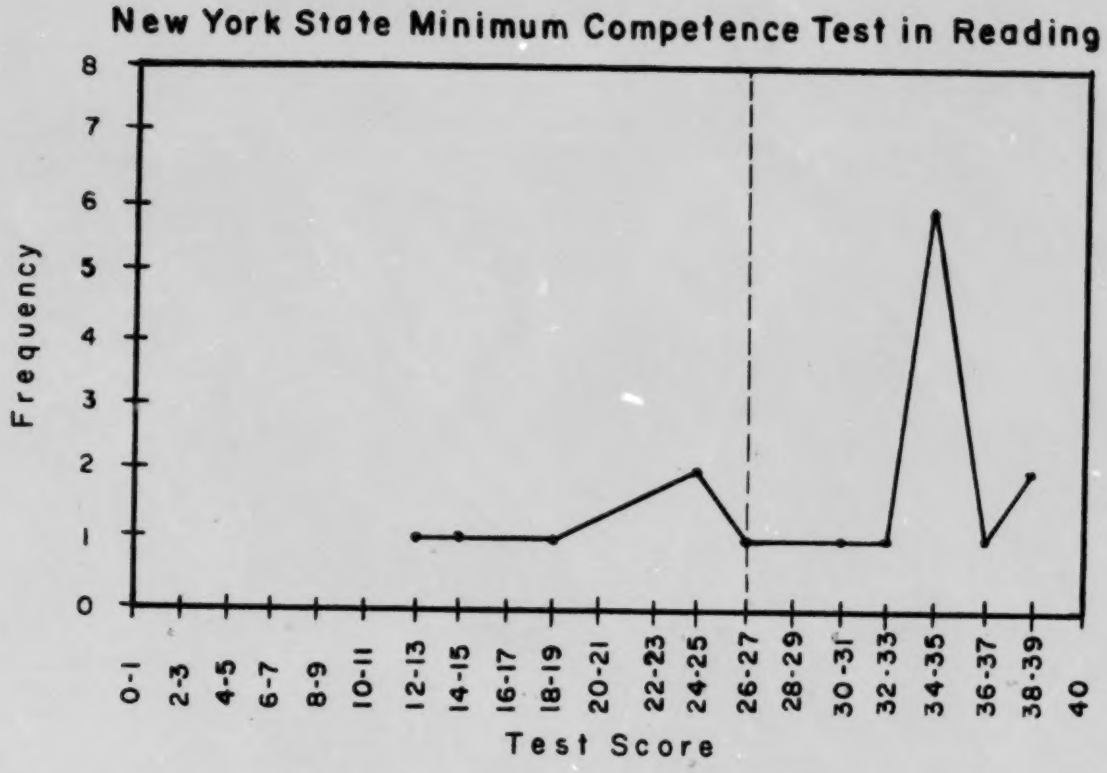


Figure 12

Frequency Polygons-Class G



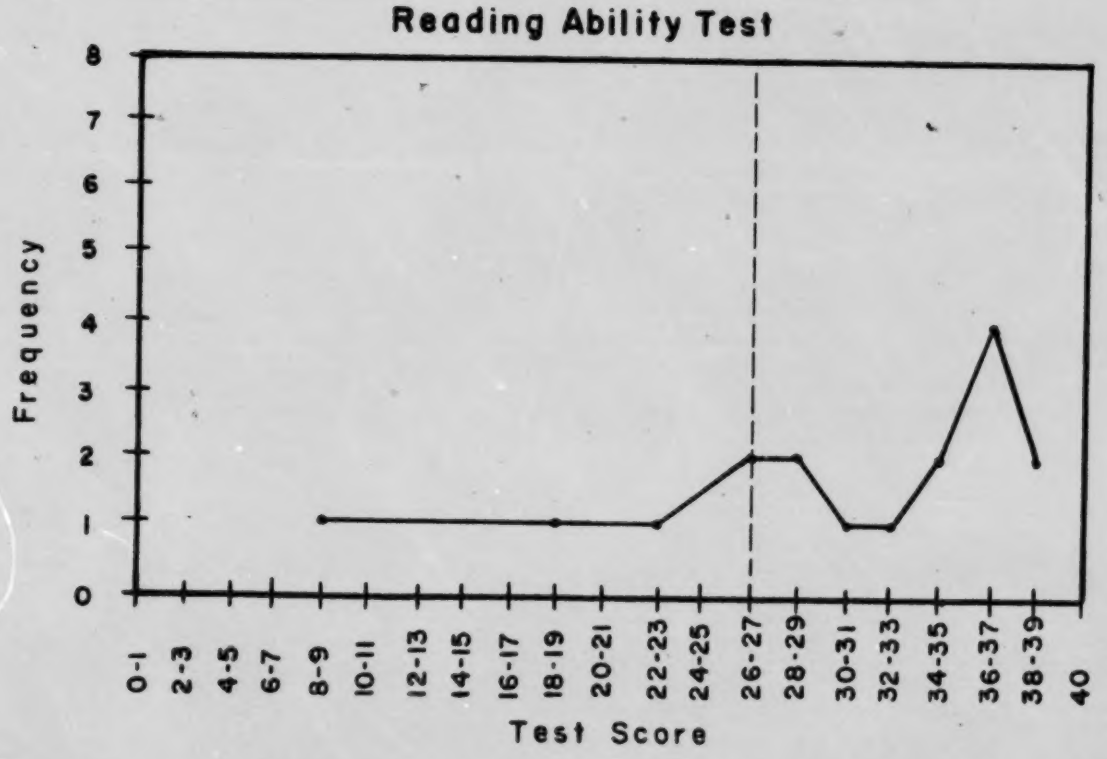
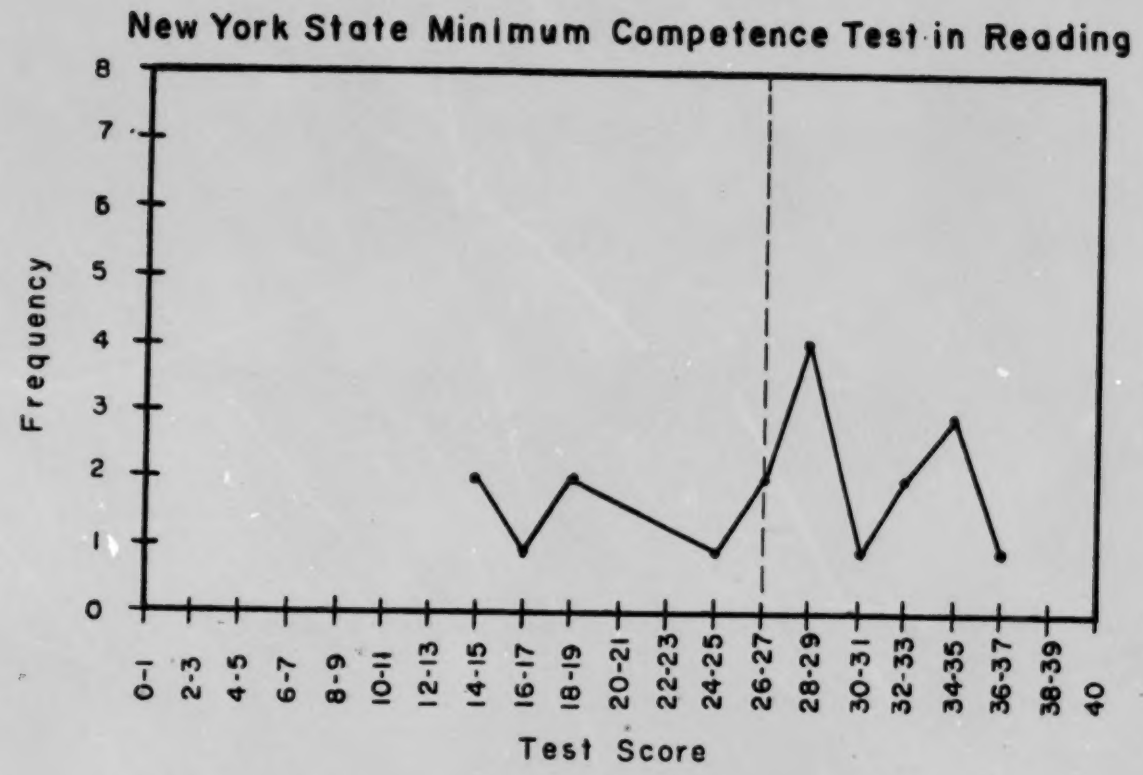
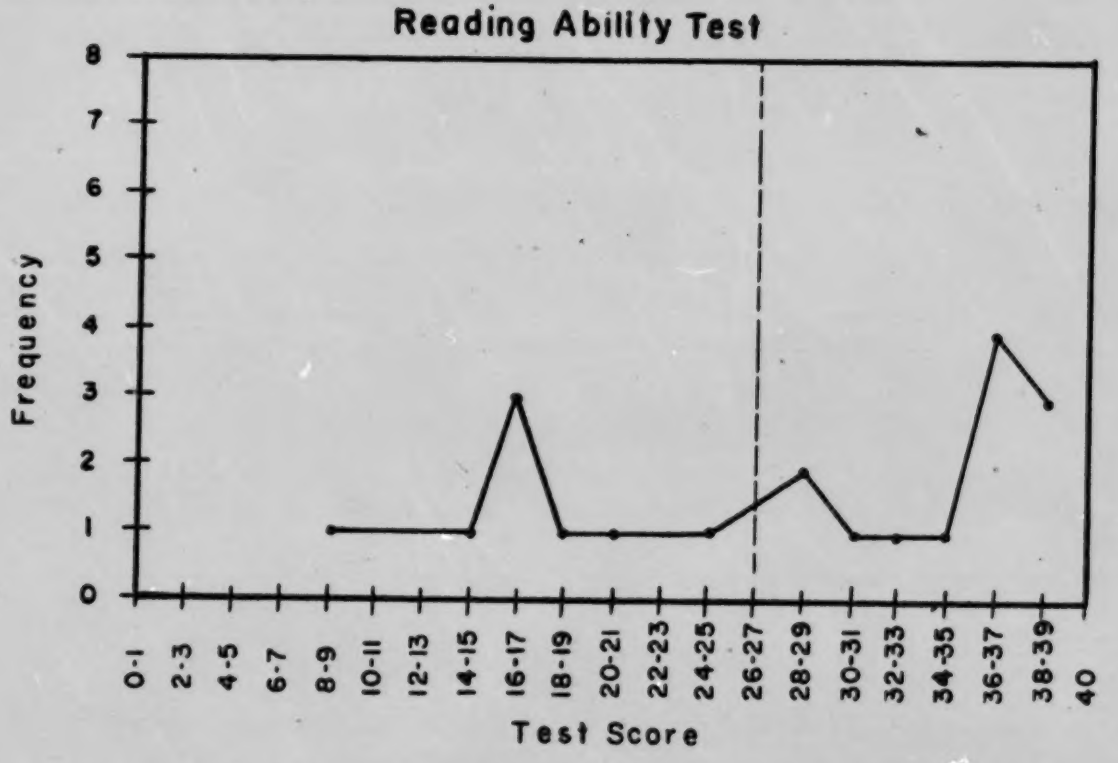


Figure 13

Frequency Polygons-Class H





H demonstrated fairly equal ability on both tests. Gains made by individual classes on the Reading Ability Test were somewhat graduated; three classes made a mean score gain of from 3 to 5 points; three classes gained 1 to 1.5 points; one class increased mean score by .6; and one class produced the same mean score on both tests. These differences were tested for significance but no significant difference was found among the classes when compared with each other. The reaction of the groups as a whole produced a level of significance but not the reaction of the sub-groups in comparison with each other.

Each class represents a similar IQ pattern according to results of the most recent Otis Quick Score Testing. The range of the total test group extends from 62 to 127 and each class is composed of the same widely varied pattern of intelligence. This probably accounts for the lack of significant difference among classes. If classes were grouped homogeneously according to intelligence a significant difference probably would have resulted in the comparison of the sub-groups.

Reliability of the Tests

Experimental conditions of the October, 1970 testing were such that they necessitated an interruption of the regular school program. Because of this, it was not possible to reschedule both tests after a suitable time lapse, again this year, in order to collect data for determination of reliability. However, in the fall of 1969, the two tests were admin-

istered on a pre-experimental basis to two classes of seventh and eighth grade children in a typical urban school in Albany. It was possible in September, 1970 to trace twenty-eight of these children and retest them with the Minimum Competence Test and the Reading Ability Test. This provided limited, but valuable, data in the determination of how reliable each measure might be especially in reference to each other.

Reliability coefficients proved unusually high with the Minimum Competence Test in Reading yielding a .90 and the Reading Ability Test yielding .93--results of both tests proved remarkably similar after a one-year period. This conclusion is based on a small sampling and coefficients would probably be reduced with a much larger experimental group. Nevertheless, coefficients produced by this sampling represent a favorable prognosis for the power of each test to measure consistently over a period of time.

Item Analysis of the Tests

An examination of the items on both the Minimum Competence Test and the Reading Ability Test was made in order to discover how testees responded to specific types of questions. Figure 14 is a graphic representation of the percentage of children who responded correctly to the forty items on each test. The solid line represents the percentage of correct responses to the Minimum Competence Test and the dotted line represents the percentage of correct responses to the Reading Ability Test.

Figure 14

Item Analysis Percentage of Correct Answers

According to the total picture, testees scored higher on twenty-two items of the Reading Ability Test, higher on sixteen items of the Minimum Competence Test and the same on two items in both tests; namely, #4 and #22.

Davis warns of the difficulty of judging the merit of test items through item analysis in the following statement:

> It is dangerous to assume that item-analysis data will necessarily reveal faults in items. The individuals included in the sample used to obtain item-analysis data are rarely experts in the fields covered by the items. Hence, the responses of the testees may, in one sense, be thought of as a pooling of ignorance.[1]

Such a danger is apparent; however, the purpose of the item analysis made in this study is not to serve as the basis for refinement of items but merely to examine how testees responded to two different types of content-related-test items. No sophisticated revision of individual items is planned.

The most obvious question raised by the percentages charted in Figure 14 is that of extreme percentages. Items falling below 60 per cent correct were examined to determine what kind of questions were involved.

On the Minimum Competence Test these included: #7, #11, #13, #20, #26, #28, #29, #36, and #37. Of these, two deal with main reason or central idea (#20, #26); four have inferences and generalizations (#7, #11, #13, #29); two

---

[1] Frederick B. Davis, Item-Analysis Data: Their Computation Interpretation, and Use in Test Construction (Cambridge, Mass.: Harvard University, 1949), p. 16.

involve vocabulary recognition (#28, #37); and one is a factual response.

On the Reading Ability Test four items fell below the 60 per cent difficulty level. These items included: an item which introduced a figure of speech (#19); two factual items (#29, #30) and a vocabulary word (#37). It is interesting to note that the same vocabulary word, "dormant," caused a problem on both tests despite the fact that it appeared in two different sets of content.

The next set of scores investigated were those in which item response was directly opposite. For example, in items #26, #28, and #36, there is a wide and opposite discrepancy, found between scores on the two tests.

Item #28 illustrates specifically how unfamiliar concepts confuse and penalize a number of urban children. The item on the Minimum Competence Test contains three concepts which are alien to urban children; namely, rennet, curd, and whey. These are words which present no serious decoding problem. It is likely that many children could pronounce these words but had no oral-aural counterpart with which to interpret them. They are literally nonsense syllables for children who have not had life-related experiences with them. Results of such an item are predictable in terms of the response of an urban child. Of the total group, only 54 per cent responded correctly to item #28 on the Minimum Competence Test; whereas, 81 per cent responded correctly to a more relevant item concerned with the differences in the effects

of pot and LSD on the Reading Ability Test counterpart item. The response of low scorers, in particular, should be noted:

Minimum Competence Test Correct Responses

Item #28	High Scorers - 79%
	Low Scorers - 29%

Reading Ability Test Correct Responses

Item #28	High Scorers - 97%
	Low Scorers - 65%

Thus, an 18 per cent increase occurred among better readers and a 36 per cent gain among poorer readers on a more relevant item which neither threatened nor confused testees by introducing foreign conceptual matter.

Item #26 presented the most dissimilar results between responses on the two tests. A 49 per cent difference occurred between responses to this item on the two tests. Only 27 per cent of the total group tested were able to select the correct answer to this item on the Minimum Competence Test; on the Reading Ability Test, 76 per cent responded correctly. Both items deal with the selection of a central theme of two paragraphs. The breakdown of responses selected in terms of the four choices for this item on each test is shown in Table 6. On the Minimum Competence Test, better readers were drawn to choice number two which emphasized the accidental discovery of cheese. This high-scoring group failed to notice that the second paragraph in the selection discussed current facts about cheese consumption. This group failed to recognize that both parts of a selection have to be encompassed by a central theme.

On the Minimum Competence Test low scorers missed the answer by the widest margin, selecting the distractors as correct responses. Only twelve managed to score correctly.

Table 6

Analysis of Item #26 in Terms of High and Low Scorers

Choices	Minimum Competence Test		Reading Ability Test	
	High Scorers	Low Scorers	High Scorers	Low Scorers
1.	4	14	1	3
2.	38	29	4	11
3.	6	17	5	11
4.	28	12	65	49
	No Response - 3		No Response - 2	

Three testees on the Minimum Competence Test and two on the Reading Ability Test failed to respond at all.

Item #36 on the two tests showed yet another distinct difference in responses. The reason for this is probably attributable to the type of concept being dealt with in each question. The Reading Ability Test posits a question which deals with a place fact whereas, the Minimum Competence Test introduces a time fact in the question. Thirty-eight per cent of the testees proved more able to cope with a place identification than with a time identification. This may be because a time fact is more of an abstraction than a place fact; the latter, thus, probably easier to interpret for most children.

Somewhat of a reversal occurs on the two tests as a whole in regard to percentage of correct answers. Figure 14, page 112, indicates that on the first half of the test items (1-20) a greater number of children scored higher percentages on the Minimum Competence Test whereas on the second half of the items (21-40) a larger number scored higher on the Reading Ability Test. Results of both tests may be seen more clearly in an analysis of the items of each test independently.

Figures 15 and 16 deal with the power of each item on each test to discriminate between high and low scoring testees. Davis points out that: "A single test item is regarded as having perfect positive discriminating power if everyone in the high-scoring group and no one in the low-scoring group answers it correctly."[1] To examine the discriminatory power of each item on both tests, Figures 15 and 16 were constructed. High-scorers are represented on the graphs by a solid line and low-scorers by a dotted line. The crucial point of discrimination is the degree to which percentage of correct responses of high- and low-scorers are opposed. The greater the distance between the high and low points, the greater the discriminatory power. Charts are based on the percentage of each item answered correctly.

An ideal discriminatory item would yield a differentiation of 100 per cent between high and low scorers. Using

---

[1]Davis, p. 1.

Figure 15

Discriminating Power of Items on the Minimum Competence Test in Reading

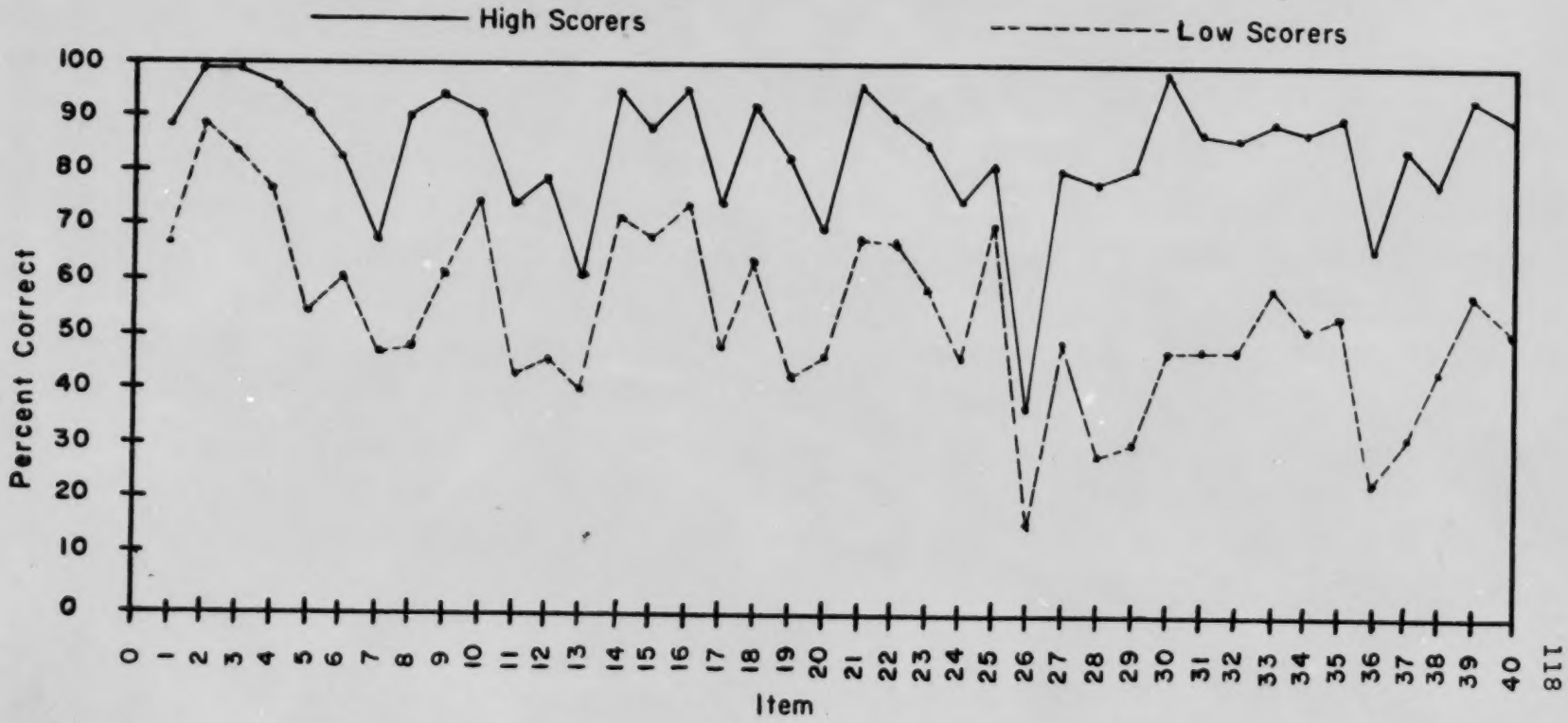
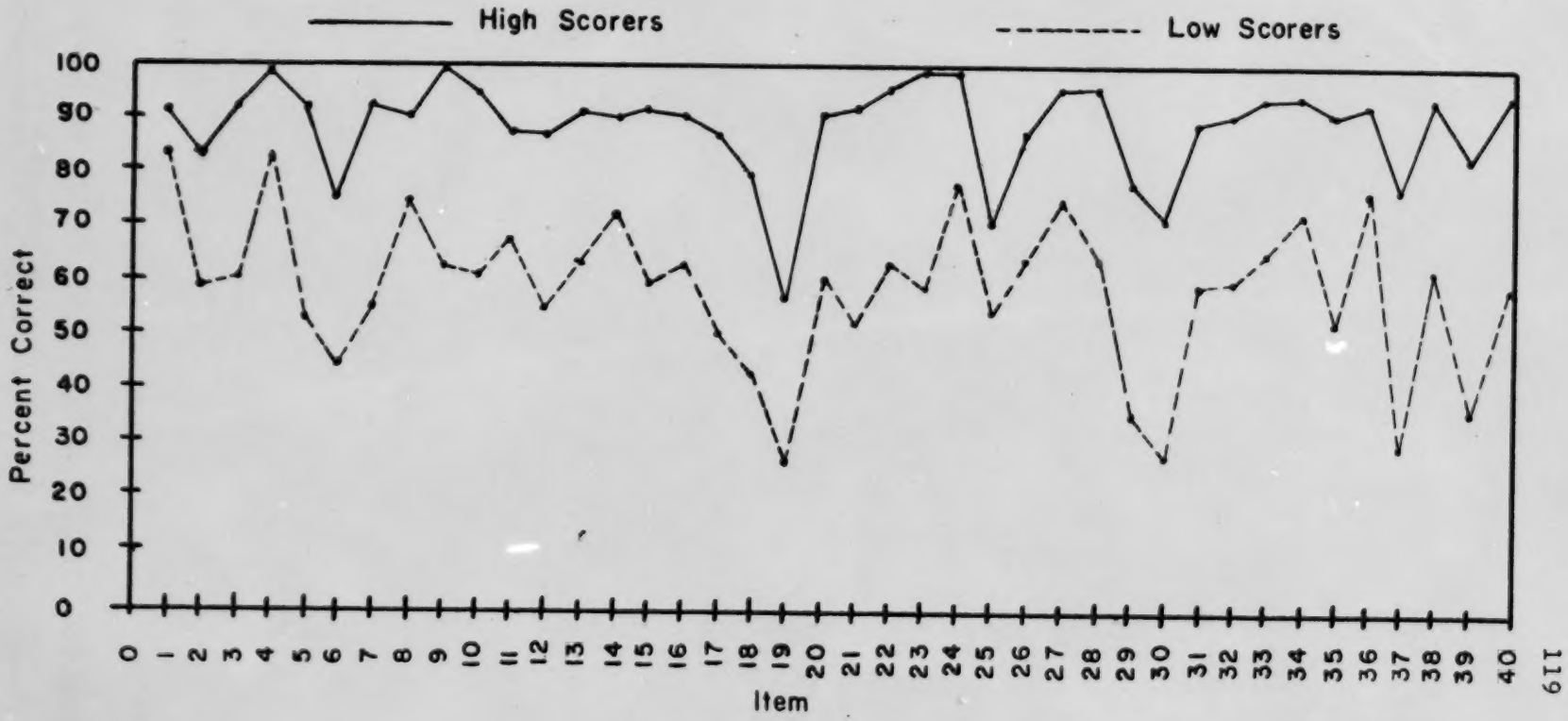


Figure 16

Discriminating Power of Items on the Reading Ability Test



this base, it is possible to designate the number of items on both tests which approach this ideal. On the Minimum Competence Test the percentage differential between high and low scorers is shown on Table 7.

Table 7

Best Discriminatory Items on the Minimum Competence Test in Terms of High and Low Scorers

Percent Differential	Items
60%+	#37
50%+	#8, 28, 29, 30
40%+	#19, 31, 32, 34, 35, 36, 39, 40
30%+	#2, 5, 11, 12, 18, 24, 27, 38

On the Reading Ability Test the best discriminatory items, according to high-low responses are shown on Table 8.

Table 8

Best Discriminatory Items on the Reading Ability Test in Terms of High and Low Scorers

Percent Differential	Items
50%+	#30, 37, 39
40%+	#5, 18, 31, 23, 29, 35, 38
30%+	#3, 6, 7, 9, 10, 12, 13, 15, 17, 19, 20, 22, 28, 31, 32, 40

Approximately half of the items on both tests may be considered good to excellent discriminators of reading ability.

The best item, in terms of discriminatory power on both tests, is #37 which creates the most substantial degree of difference between the response of high- and low-scorers. That item is concerned with use of the vocabulary word, "dormant." Incorporation of this unfamiliar word which is meant to test contextual ability resulted in the fact that better readers are more adept at determining meaning of a word from context than less capable readers most of whom failed on this item.

It is important to note that although both tests share a similar degree of discriminatory power, the pattern of discrimination in the two tests is noticeably different. In Figure 15, page 118, for example, the pattern of discrimination undergoes a decided change between the beginning and the end of the test. A fairly parallel pattern of answers is established between item #1 and item #25. Item #26 creates a marked turning-point after which the pattern changes completely. Both high and low scorers found item #26 to be the hardest on the Minimum Competence Test. This item deals with the central theme of a selection which is concerned with the history of cheese and its current consumption in the United States. According to the Dale-Chall Reliability Formula, recorded on page 76 of Chapter 3, paragraph I related to this item tested out at the eleventh-twelfth grade of difficulty. The new pattern which developed among low scorers

is attributed to the fact that poorer readers began to experience both fatigue and frustration by the time they reached paragraph I. They collapsed at item #27, the first question related to paragraph I, and continued in a relatively downward trend.

The pattern established by results of the Reading Ability Test is much more parallel throughout. Peaks and valleys in Figure 16, page 119, are neither as frequent nor as sharp. High scorers sustain a more consistent line indicating that they are more favorably affected by relevant content. Low scorers, as well, follow a pattern which, although marked by several sharp drops is somewhat less erratic than on the Minimum Competence Test. It is important to note that where low-scorers did drop significantly on both tests, items were related to paragraphs which were written above ninth grade level. Examples of this on the Minimum Competence Test are: #19, #26, #28, #37; examples on the Reading Ability Test include: #19, #30, #37, #39.

Figure 16 demonstrates the more sustained pattern of ability which was achieved by the total experimental group on a test of more relevant reading material. A fatigue and frustration factor appeared in Figure 15; however, the incorporation of high interest reading content appears to have prevented this from occurring on the Reading Ability Test. Foreign content appears to cause more unpredictable results.

VanDalen and Meyer warn against the danger of over-analyzing data and caution that when simple statistics answer questions, "there is little merit in performing more complicated analysis for the sake of window dressing. . . ;"[1] this study now borders on over-analysis. Sufficient statistics have been considered all of which compliment each other and all of which point to the fact that the nature of reading content does indeed influence reading ability. Results considered support rejection of the null hypothesis and confirm the suspicion that a number of children in urban areas are penalized unnecessarily by the alien nature of reading matter presented on at least one typical test of reading.

---

[1]Deobold B. VanDalen and William J. Meyer, Understanding Educational Research (New York: McGraw-Hill Book Company, 1966), p. 330.

Chapter 5

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

SUMMARY

Problem

A marked discrepancy has been noted in Albany in the course of the past five years. Classroom evidence points to the fact that urban children are reading better every year but annual test results fail to reflect this observable pattern of growth. This study was designed to investigate if the tools of measurement are, in any way, deflecting authentic growth patterns.

One test of reading, the New York State Minimum Competence Test in Reading was chosen for investigation. It was felt that the nature of the reading content was alien to the realistic background of urban children and, as such, penalized them unfairly.

In order to test the impact of unfamiliar content on urban children, it was necessary to construct a new test. The new test, entitled the Reading Ability Test, was written to be congruent with the Minimum Competence Test. It shares with its counterpart similar grammatical structure, flow of thought, pattern of questioning, external format, and level of difficulty. The latter was established by applying the

Dale-Chall Formula of Readability to each paragraph of both tests. By holding these factors constant and substituting within the same framework a new body of relevant content, it was possible to collect data which would test the impact of content on reading achievement.

A null hypothesis was formulated which stated that: on two equally difficult but content differentiated tests of reading ability there would be no significant difference in the scores of a group of ninth grade children in an urban area. It was anticipated that it would be possible to reject the hypothesis on the basis of evidence produced. Rejection would be based on the degree to which the tests produced dissimilar results. A .05 level of confidence was stipulated as the point of rejection. The extent of differentiation would be attributed to the nature of the content in each test since other factors were controlled.

Procedure

The Minimum Competence Test and the Reading Ability Test were administered to the ninth grade class of 151 pupils at Livingston Junior High School on October 8, 1970. One half of the group took the Minimum Competence Test in the first part of the test session; the other half took the Reading Ability Test first. Testing was reversed in the second part of the test session. Test papers were scored and results analyzed for total group differences, sub-group differences, and both inter- and intra- item analysis.

Results

A significant difference was found between mean scores of the two tests. The z score produced was 2.3 and although slightly below the .01 level was well within the .05 level of confidence.

Computation of the coefficient of alienation resulted in t score of 12.9 which is significant at the .01 level of confidence. This highly significant level indicates the pronounced difference in the response of the group to two distinct types of content.

Mean score differences of sub-groups were compared. Both blacks and whites and boys and girls showed gains on more relevant material but no substantial difference appeared between the performance of the sub-groups when compared with each other. White boys and black girls made somewhat greater gains than black boys and white girls; all groups made positive gains on the Reading Ability Test as evidenced by increased mean scores.

Individual classes were compared for differences. Here, too, although a marked trend toward improved achievement appeared in six out of the eight classes, no highly significant difference could be established for one or more classes over any of the others. The broad range of intellectual ability found in each class most likely accounted for the similarity of response within classes.

Comparison of the two tests through item analysis produced a reasonable distinction between the high and low scor-

ing groups. The spread across available alternatives was in the opposite direction indicating that the items on both tests were working in the proper manner. There was also evidence of a reasonable array of difficulty. The pattern of discriminatory items was somewhat different on each test. The Minimum Competence Test in Reading gave evidence of less discriminating items at the beginning of the test and better discriminators toward the end. The Reading Ability Test, on the other hand, showed a more consistent spread of discrimination throughout.

Both tests reflected sharp drops of both high and low scorers on paragraphs which tested out above ninth grade difficulty.

A reliability coefficient was determined on a limited sampling of seventh and eighth grade children who took both tests one year apart. Coefficients of .90 and .93 were established for the Minimum Competence Test and the Reading Ability Test respectively. Both tests appear to yield highly similar results over an extended period of time.

CONCLUSIONS

The analysis of data just considered seems to warrant the following conclusions:

1. More relevant reading content on tests of minimal reading ability produces better results among a number of urban children.

2. The nature of reading content is crucial in the

determination of a minimal competence reading level. Concepts which are unfamiliar to urban children tend to deflect results in a negative direction and fail to produce accurate measurement.

3. The New York State Minimum Competence Test in Reading is invalid for those children who can read but cannot simultaneously deal with what, for them, constitutes alien reading matter.

4. In this experiment, no significant difference was found to exist among sub-groups; boys did not score better than girls; blacks did not score better than whites; nor did any one class of the eight classes tested significantly surpass any other.

5. The influence which the nature of content exerts on reading ability found in this experiment may exist in other tests of reading used at other grade levels.

6. If, as demonstrated in this experiment, a number of children can read with competence within, but not beyond, their experiential range, many more children than is currently believed may actually possess minimum competence in reading.

7. Although developing reading skill in order to benefit from vicarious experiences is a major goal of education, it would seem that the test of minimum competence is an inappropriate place in which to fulfill this goal.

8. A number of children are currently labelled as non-competent in reading who actually possess authentic mini-

mal reading ability. These children fall into three categories:

a. Children who can read and comprehend material written within their experiential background and material which considers their realistic conceptual development.

b. Children who read accurately and precisely but who read so slowly that they cannot fulfill the minimal requirements within a specified time-span.

c. Children who, on the Minimum Competence Test in Reading, answered questions correctly on material written at or below their grade level but not above it.

9. The inaccurate measurement which alien content produced in relation to a reading test may also be true of measures of intelligence. Many more types of tests at many other levels deserve similar examination.

10. Emphasis on the humanization of the curriculum carries within it the implication that tests of reading must be constructed consistent with the realistic concept-development-process of the child.

RECOMMENDATIONS

There is an obvious need for replication of this study in other urban areas in order to determine if the trend established between reading content and comprehension is universal.

A crucial question raised by this study is the definition of minimum competence in reading itself. "Is a child competent in reading if he can read and understand material written within the realistic framework of his experience?" Tests of reading often demand that he read beyond this range. Does coupling the ability to decode with the ability to respond to alien conceptual reading matter penalize urban children unfairly? Answers to these questions cannot be made without considering what competence in reading implies among sophisticated readers. Must the mathematician, for example, be able to interpret lyric poetry to be truly competent in reading? Need a linguist be able to interpret a musical score to read competently? The ramifications, although exaggerated, underscore the basic question of just what reading functions children must be able to perform in order to be considered competent in the skill.

This study was not undertaken as a means of lowering reading standards. Its purpose was to investigate if children are being measured validly in terms of the content presented to them on typical tests of reading. A trend has been established toward better achievement on relevant content. Presence of that trend deserves to be examined on other tests of reading and at other grade levels.

Another area in which additional research is necessary is that of concept development as it relates to the testing of reading skills. An assumption which appears to be true of typical tests of reading is that the mechanical linkage

between words and objects developed through years of schooling makes any area of curriculum a sufficient base on which to construct test items. What this assumption ignores is the fact that concept development is vitally related to processes and not merely verbal associations. It would seem that test construction should closely parallel concept development and relate more directly to processes rather than merely to verbal associations.

In conclusion, this study proposes that there is an immediate and urgent need for additional research which will clarify the accuracy with which children, particularly children in urban areas, are being measured in reading. It is recommended that the national "Right to Read" Program currently being advocated throughout the country include one additional right; and that is: the right of every child to have his authentic reading ability measured accurately and validly.

Recent national statistics concerning the reading status of American children have produced grim results. The reading issue is currently the most publicized failure in education.[1]

In New York State the Commissioner of Education[2] has recently revealed the shocking number of children who are

---

[1]National School Public Relations Association, Reading Crisis: The Problem and Suggested Solutions (Washington: National School Public Relations Association, 1970), p. 1.

[2]Times Union (Albany), January 15, 1971, p. 23.

considered as reading below minimum competence at this time.

Would it not be strange, if the tools of measurement in reading, rather than the children being measured, were in large part responsible for such bleak national and state reading test profiles?

BIBLIOGRAPHY

BIBLIOGRAPHY

BOOKS

Aaron, Ira E. Teaching Comprehension Skills. Atlanta, Ga.: State Education Department, 1964.

Baratz, Joan C. and Roger W. Shuy (eds.). Teaching Black Children to Read. Washington: Center for Applied Linguistics, 1969.

Bibby, Cyril. Race, Prejudice, and Education. New York: Frederick A. Praeger, 1960.

Bamman, Henry A., Midori Hiyama, and Delbert Prescott. World of Ideas: A Guide to Effective Reading. Berkeley Heights, N. J.: Field Educational Publications, Inc., 1970.

Boelen, Bernard J. Existential Thinking: A Philosophical Orientation. Pittsburgh: Duquesne University Press, 1968.

Carroll, J. B. Language and Thought. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Center for Applied Linguistics. Non-Standard Speech and the Teaching of English. Washington: Center for Applied Linguistics, 1964.

Coburn, Wilbur. Research and Development: Advance in Education. Washington: U. S. Government Printing Office, 1968.

Cohen, S. Alan. Teach Them All to Read: Theory, Methods, and Materials for Teaching the Disadvantaged. New York: Random House, 1969.

Crary, Ryland. Humanizing the Curriculum. New York: Harcourt Brace and World, 1970.

Davis, Frederick B. Item-Analysis Data: Their Computation, Interpretation and Use in Test Construction. Cambridge, Mass.: Graduate School of Harvard University, 1949.

Dechant, Emerald. Diagnosis and Remediation of Reading Disability. West Nyack, N. Y.: Parker Publishing Company, 1968.

Dunning, Stephen (ed.). English for the Junior High School Years. Chicago: National Council of the Teachers of English, 1969.

Fitzgibbon, Thomas J. (ed.). Evaluation in the Inner City, Report of an Invitational Conference on Measurement in Education. New York: Harcourt, Brace and World, 1970.

Fries, Charles C. Linguistics and Reading. New York: Holt, Rinehart and Winston, 1966.

Gans, Roma. Common Sense in Teaching Reading. New York: The Bobbs Merrill Company, 1963.

Gage, N. L. (ed.). Handbook of Research on Teaching. Chicago: Rand McNally and Company, 1963.

Getzels, Jacob.W. Developing Permanent Interest in Reading. Chicago: University Press of Chicago, 1956.

Good, Carter V. Introduction to Educational Research. New York: Appleton-Century-Crofts, 1963.

Guilford, J. P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill Book Company, 1956.

Havighurst, Robert J. and Bernice L. Neugarten. Society and Education. Boston: Allyn and Bacon, Inc., 1962.

Helmstadter, G. C. Principles of Psychological Measurement. New York: Appleton-Century-Crofts, 1964.

Hoover, Kenneth H. Learning and Teaching in the Secondary Schools. Boston: Allyn and Bacon Inc., 1968.

Horricks, John E. Assessment of Behavior: The Methodology and Content of Psychological Measurement. Columbus, O.: Charles E. Merrill Books, Inc., 1964.

Jersild, Arthur T. Child Psychology. Englewood Cliffs, N. J.: Prentice-Hall Inc., 1968.

Koenker, Robert H. Simplified Statistics. Chicago: McKnight and McKnight Publishing Company, 1961.

Kohl, Herbert. 36 Children. New York: The New American Library Press, 1967.

Kolesnik, Walter B. Educational Psychology. New York: McGraw-Hill Book Company, Inc., 1963).

Kwant, Remy C. Phenomenology of Language. Pittsburgh: Duquesne University Press, 1965.

Lannon, Roger T. (ed.). Evaluation in the Inner City. New York: Harcourt, Brace and World, 1970.

Leacock, Eleanor Burke. Teaching and Learning in City Schools, a Comparative Study. New York: Basic Books, Inc., 1969.

Lee, Dorris and Roach Van Allen. Learning to Read Through Experience. New York: Appleton-Century-Crofts, 1963.

Lerner, Daniel. The Passing of Traditional Society: Modernizing the Middle East. New York: Free Press, 1958.

Loretan, Joseph O. and Shelly Umans. Teaching the Disadvantaged. New York: Columbia University, Teachers College Press, 1966.

Lowry, Ritchie P. and Robert P. Rankin. Sociology, The Science of Society. New York: Charles E. Scribner's Sons, 1969.

May, Rollo. Man's Search for Himself. New York: W. W. Norton and Company, Inc., 1967.

Miller, Harry L. Education for the Disadvantaged. New York: The Free Press, 1967.

Morgan, Clifford T. and Richard A. King. Introduction to Psychology. New York: McGraw-Hill Book Company, 1966.

National Reading Conference, Inc. Reading: Process and Pedagogy. Nineteenth Yearbook of the National Reading Conference. Milwaukee, Wisc.: The National Reading Conference, Inc., 1971.

National School Public Relations Association. Reading Crisis: The Problem and Suggested Solutions. Washington: National School Public Relations Association, 1970.

National Society for the Study of Education. Development In and Through Reading. Sixtieth Yearbook, Part I. Chicago: University of Chicago Press, 1961.

Raths, Louis E., Merrill Harmin, and Sidney B. Simon. Values and Teaching: Working with Values in the Classroom. Columbus, O.: Charles E. Merrill Publishing Company, 1966.

Reading Reform Foundation. Proceedings of the Eighth Annual Conference. San Francisco: Reading Reform Foundation Press, 1969.

Rees, Helen E. Deprivation and Compensatory Education. New York: Houghton Mifflin Company, 1968.

Rogers, Everett M. Modernization Among Peasants: The Impact of Communications. New York: Holt, Rinehart and Winston, 1969.

Schick, George B. A Guidebook for the Teaching of Reading. Chicago: Psychotechnics Press, Inc., 1966.

Smith, Henry Lee. English Morphophonics: Implications for the Teaching of Literacy. New York: New York State English Council, 1968.

Spache, George D. Good Reading for Poor Readers. Chicago: Garrard Publishing Company, 1966.

Stauffer, Russell. The Language-Experience Approach to the Teaching of Reading. New York: Harper and Row, Publishers, 1970.

Thorndike, Robert L. and Elizabeth Hagen. Measurement and Evaluation in Psychology and Education. New York: John Wiley and Sons, Inc., 1961.

University of the State of New York. Closing the Gap: A Report of the First Two Years of Experience With Title I in New York State. Albany: State Education Department, 1968.

_____. Manual of Directions for the Minimum Competence Test in Reading. Albany: State Education Department, 1969.

_____. Reading in Secondary Schools. Albany: State Education Department, 1965.

VanDalen, Deobold and William J. Meyer. Understanding Educational Research. New York: McGraw-Hill Book Company, 1966.

Vygotsky, Lev Semenovich. Thought and Language. Translated by Eugenia Hanfmann and Gertrude Vakar. Cambridge, Mass.: Massachusetts Institute of Technology, 1962.

Warner, Sylvia Ashton. The Teacher. New York: Simon and Schuster, 1963.

Whipple, Gertrude and Millar Black. Reading for Children Without, Our Disadvantaged Youth. Newark, Del.: International Reading Association, 1966.

PERIODICALS

Behavior Today: The Human Sciences Newsletter, I, 7 (May 25, 1970).

The Times Union (Albany, N. Y.), January 15, 1971, p. 3.

Dale, Edgar and Jeanne S. Chall. "A Formula for Predicting Readability," Educational Research Bulletin, XXVII, 1, 2 (1948), 1-9, 11-28.

Dawkins, John. "A Reconsideration of the Dale-Chall Formula," Elementary English, XXXIII (December, 1956), 515-519.

Ebel, Robert L. "The Curriculum and Achievement Testing," Educational Technology, V (May, 1970), 22-24.

Erickson, Frederick D. "'F'get You Honky!': A New Look at Black Dialect and the School," Elementary English, XXXXVI, 4 (April, 1969), 494-499, 516-517.

Finder, Morris. "Comprehension: An Analysis of the Task," Journal of Reading, XIII (December, 1969), 199-202.

Frymier, Jack R. "Motivating Students to Learn," NEA Journal (February, 1968), 37-39.

Glaser, Robert. "Instructional Technology and the Measurement of Learning Outcomes: Some Questions," American Psychologist, XVIII (August, 1963), 519-521.

Hall, Elizabeth, "A Conversation With Jean Paiget," Psychology Today, III (May, 1970), 25-32.

Hardman, Laurence L. "Slow Readers - A Happy Experience," English Journal, LVII, 3 (March, 1968), 405-408.

Jensen, A. R. "How Much Can We Boost I.Q. and Scholastic Achievement?" Harvard Educational Review, XXIX (1969), 1-123.

Klein, Stephen. "Evaluating Tests in Terms of the Information They Provide," Evaluation Comment, II,2 (June, 1970), 1-4.

Klapp, Orrin E. "Heroes, Villains and Fools As Agents of Social Control," American Sociological Review, IXX (February, 1954), 56-62.

National Center for Educational Statistics. "Advance Statistics for Management," V (April, 1970), 1-2.

Ozman, Howard A. "Value Implications in Children's Reading Material," The Reading Teacher, XXII (December, 1968), 246-250.

Popham, James W. and T. R. Husek, "Implications of Criterion-Referenced Measurement," Journal of Educational Measurement, VI, 1 (Spring, 1969), 1-9.

Rosenthal, Robert, "Self-Fulfilling Prophecy," Psychology Today, II (September, 1968), 46-51.

Smith, Nila B. "Literature for Space-Age Children," Education Magazine, October, 1960, p. 5.

Smith, Richard J. and Karl D. Hesse, "The Effects of Pre-reading Assistance on the Comprehension of Good and Poor Readers," Research in the Teaching of English, III, 2 (1969), 166-177.

Skager, Rodney W. "Objective Based Evaluation: Macro-Evaluation," Evaluation Comment, II, 2 (1970), 1-4.

Woodring, Paul (ed.), "Education in America," Saturday Review, July 18, 1970, p. 41.

OTHER SOURCES

Dissertations

Bentall, Grace. "Determination of Critical Levels of Reading Ability for High School Students as Measured by Course Marks and Achievement Test Results." Doctoral Dissertation, University of Oregon, 1961.

Einstein, Florence. "A Comparative Study of Reading Achievement of Seventh-Grade Junior High School Boys in a Low Socio-Economic Area of New York City." Doctoral Dissertation, New York University, 1960.

Storey, A. G. "A Reading Comprehension Test for Junior High School Pupils in Alberta." Doctoral Dissertation, Stanford University, 1960.

Ware, Ralph H. "The Development of a Criterion-Related Reading Comprehension Test for Exposition and Argument." Unpublished Doctor's Dissertation, State University of New York at Albany, 1970.

Unpublished Report

Sister Mildred, "Reading Comes Alive: Instructional Materials Center Keeps It Breathing," report submitted to the Albany District Title I Office, 1969.

Scholarly Papers

Pittenger, Owen. "Psychological Considerations in Identifying Disadvantaged Readers." Paper read at Reading Symposium, State University of New York at Oswego, August 12, 1969.

Wardeberge, Helen. "Reading: How to Develop Skills and Motivation." Paper read at Sperry High School, Rush-Henrietta School District, March 28, 1969, New York.

APPENDICES

APPENDIX A

MINIMUM COMPETENCE TEST IN READING
for New York State Secondary Schools
Experimental Edition

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO.

This is a test to see how well you understand what you read. It contains 11 reading passages. You are to read each passage and answer the questions based on it. The correct answer for each question will be found among the choices following the question. Decide which choice is the correct answer. Then, on the answer sheet, blacken the answer that has the same number as the correct answer. You will have 40 minutes to answer the 40 questions.

To make sure that you know what to do, read the sample passage below and see how the questions are answered:

SAMPLE READING PASSAGE

Wood has many uses. Just take a look at the things around you. Wood is seen on all sides. Doors, pencils, chairs, desks, bookshelves, rulers and a great many other things are made from wood. Even the paper on which this is written is made from woodpulp.

1. The story is mainly

1. building materials
2. woodpulp
3. uses of wood
4. kinds of wood

The correct answer is answer number 3. Therefore, a heavy black line has been drawn in the answer space under 3 on the sample answer sheet.

SAMPLE ANSWER SHEET

	1	2	3	4
1			■	
	1	2	3	4
2		■		

2. Woodpulp is used to make

1. chairs
2. paper
3. rulers
4. doors

The correct answer is answer number 2. Therefore, a heavy black line has been drawn in the answer space under 2 on the sample answer sheet.

In the same way, on your answer sheet, answer all the questions on the reading passages. You will find the questions for each reading passage below that passage. Remember, all answers must be marked on the separate answer sheet. Blacken only one answer space for each question. Work rapidly and do the best you can.

DO NOT MAKE MARKS OF ANY KIND IN THE TEST BOOKLET.

All Rights Reserved
Unauthorized reproduction or use prohibited

The University of the State of New York
The State Education Department, Albany 1

DIRECTIONS: Read each story and answer the questions.
Mark your answers on the separate answer sheet.

Every kind of animal occasionally produces offspring that are white in color. Even white lions, lobsters and giraffes can be seen if you visit the right place at the right time. All these creatures are called albinos. Some human beings are born with a milky-white skin and pink eyes. They also are known as albinos. Whole tribes of albinos have been found in Africa. The human albino is very sensitive to sunlight and sunburns very easily. The albino is not really so strange as he seems. Actually, he merely lacks the pigment which gives the rest of us our coloring.

1. The story says that albinos are

   1. plants
   2. white in color
   3. natives of Africa
   4. mostly human

2. Albinos must guard against

   1. Africans
   2. lions
   3. cold weather
   4. the summer sun

3. In this story, the word "pigment" means

   1. albino pigs
   2. pigs and men
   3. coloring matter
   4. strange colors

The British Isles are not much larger than the state of Minnesota. And in that small space there are more than 50,000,000 people, which is nearly one-third of the total population of the United States. Living under such crowded conditions naturally gives most Englishmen a deep desire to get away from each other. And so, where the American has a front porch in order to keep an eye on what his neighbors are doing, the Englishman has a back porch in order to prevent his neighbor from seeing what he is doing.

4. According to the story, living conditions in the British Isles are

   1. crowded
   2. unpleasant
   3. cheerful
   4. good

5. The story indicates that Englishmen are

   1. more neighborly than Americans
   2. interested in knowing what their neighbors are doing
   3. anxious to get away from the British Isles
   4. interested in getting away from each other

6. The story says that

   1. Englishmen like back porches better than front porches
   2. Americans like back porches better than front porches
   3. there are no front porches in the British Isles
   4. there are no back porches in America

Riding among the first crude settlements of Indiana and Missouri, John Mason Peck preached the Gospel to any who would stop their work. With his Bible propped up on the saddle in front of him, he would read and make up sermons as he rode from place to place. Yet his concern for religion did not keep him from noticing the manners, customs, clothes, habits and ways of life of the people he met. He found out that many pioneers had started west with no knowledge of the problems they would face, and that few of them had equipment capable of coping with the wilderness. He decided it was his duty to help other people who were planning to become pioneers, and so he wrote a book called Guide for Emigrants to the West.

7. The story says that many of the pioneers whom Mr. Peck met were

1. poorly prepared to succeed in the wilderness
2. religious people who put their faith in God rather than in their equipment
3. ignorant people who could not read the Bible for themselves
4. natives of Indiana and Missouri

8. The phrase "coping with" is closest in meaning to

1. discovering
2. handling
3. solving
4. guiding

9. Mr. Peck wrote a book because he wanted to

1. tell people what the pioneers were like
2. help those who were getting ready to travel into the West
3. make money to give to the pioneers who were suffering hardships
4. spread the Gospel among the settlements he could not reach on horseback

There are places in the world where water is so scarce that it is almost worshipped. But whenever anything is easy to get, people usually treat it carelessly. They forget how valuable it may be. In a land like ours - with winter snows and summer rains, with rivers and brooks and lakes and ponds - we don't think much about water. Once in a while we are concerned when we have too much and there may be a flood, or too little and the crops may dry up. Least concerned of all is the city dweller. All he usually has to do is turn on the tap to get as much as he needs for drinking and washing.

10. The writer is trying to

    1. remind us that water is more important than we seem to realize
    2. make fun of city dwellers, who think water naturally flows through pipes
    3. tell us how foolish those people are who worship water
    4. warn us of the dangers connected with our excellent water supply

11. According to the writer, he lives in a land where there

    1. is much rain and little snow
    2. are frequent dry spells
    3. are seasons of snow and seasons of rain
    4. are long, hard winters and short, hot summers

12. The writer indicates that people who experience occasional floods

    1. worship water in other lands
    2. forget that crops dry up
    3. turn the tap on carelessly
    4. are reminded of water problems

13. The writer is telling us that

    1. people usually appreciate their blessings
    2. people value most of nature's gifts, but take some for granted as being without value
    3. when we get things with little effort, we do not appreciate them as we should
    4. what we get for nothing is usually worth no more than we have paid for it

For more years than anyone can remember, the American ideal of a proper home has been a house with a yard around it. Now, around the country, a different kind of a house is popping up. These new houses are row houses. Each has its own front, its own back, but shares side walls with its neighbors. These houses are being built because they cost less. But the biggest saving comes in land cost, as less land is required for each house. 146

Actually, row houses have been with us since before the Revolution. However, the reputation of the row house has not been enviable. Too many row houses have become slums and many people have fled to the suburbs to escape them.

14. A house in the middle of a row will have its own

1. front and one side
2. back and one side
3. four sides
4. front and back

15. If you purchase a row house, your biggest saving will be in the cost of the

1. building
2. land
3. heating
4. yard upkeep

16. Row houses declined in popularity for some time because they were

1. often found in slum areas
2. too expensive to build
3. unsuitable for city lots
4. too small for average families

Mark Twain once said that everybody talks about the weather, but nobody does much about it. Scientists tell us that the day may come when this saying is no longer true. But until that time, we should know the weather and understand it.

Studies have shown that weather has an amazing effect on our moods, disposition, and even creativity. For example, should the barometer start dropping steadily, you're likely to feel restless, irritable and pessimistic, while a steady rise in barometric pressure is apt to make you more relaxed, satisfied with life in general, and less likely to have an accident.

As in the case of many other occurrences on the planet earth, weather in all its forms, from a breeze to a tornado, can be traced to Old Sol. Of course the atmosphere is involved too, but the sun is the master chef stirring up all the ingredients for weather and climate as we know it.

17. This story may be described as

1. biographical
2. informational
3. fictional
4. humorous

18. According to the story, when the barometer falls steadily, people are likely to feel

1. gay
2. creative
3. optimistic
4. depressed

19. The words "Old Sol" (last paragraph) are used to refer to

1. Mark Twain
2. scientist
3. the sun
4. the earth

Michelangelo is probably the greatest figure in the history of architecture who did not consider himself primarily an architect. Only Phidias can rival him as a sculptor turned builder. Michelangelo's buildings, such as St. Peter's, really belong to no architectural tradition at all. They are in fact gigantic works of sculpture in which many previous traditions of historical architecture were transformed by his unique power. The work of Michelangelo is a creation, not a revision or improvement of traditional styles. He managed to convey, in architecture as in sculpture, his view of the human body as an earthly prison that confines the flight of the soul.

147

20. The main reason why Michelangelo was better than other architects is that he

1. copied old styles
2. paid attention to the little things
3. refused to build like a sculptor
4. created new ideas in building

21. The story says that, before Michelangelo was an architect, he was a

1. sculptor
2. religious man
3. poet
4. stonemason

22. As used in this story, the word "managed" means

1. liked
2. gave orders
3. studied
4. was able

The first newspaper in the American colonies was printed in 1690 by Benjamin Harris. It was 3 pages long and 6 inches by 9½ inches when folded. Four days after it came off the press, a law that made printing it illegal was passed. The governor was afraid that something against the government might be put into print. Today, freedom of the press is important to American democracy.

23. The size of the first newspaper was

1. about the size of a dollar bill
2. much smaller than the newspaper of today
3. the same size as the newspaper of today
4. much larger than the newspaper of today

24. The colonial governor was afraid because Harris' paper

1. might say he was a poor governor
2. would very likely print lies
3. was breaking the law
4. was a threat to American democracy

25. The story says that the first newspaper was

1. written by hand
2. written in blue ink
3. printed on a machine
4. printed on a typewriter

Cheese is a very important dairy product. No one knows who made the first cheese, but according to an ancient legend, it was first made accidentally by an Arabian merchant. The merchant put his supply of milk into a pouch made of a sheep's stomach when he set out on a long day's journey across the desert. The rennet in the lining of the pouch combined with the heat of the sun to cause the milk to separate into curd and whey. He found at nightfall that the whey satisfied his thirst and the cheese satisfied his hunger. 148

Today most cheese is made from cow's milk, simply because cows are milked more generally throughout the world than other animals. People all over the world like and eat cheese. People in the United States are no exception. Although we do not eat nearly so much cheese as do people in some countries, we eat nearly twice as much now as 20 years ago. More than a billion pounds of cheese is now made in the United States each year. About one-tenth of our annual milk production is used in making this cheese.

26. The central theme of this story deals with

1. how to make useful products from milk
2. an accidental discovery of a nutritious food
3. the importance of the cheese industry in the United States
4. information about a leading dairy product

27. When milk is separated, the part that forms cheese is called

1. curd 2. rennet 3. whey 4. yeast

28. According to the story, which would always be found in a sheep's stomach?

1. milk 2. rennet 3. curd 4. whey

29. According to the story, cow's milk is usually used in making cheese because

1. satisfactory cheese cannot be made from other kinds of milk
2. cow's milk is easier to process than other kinds of milk
3. cow's milk is more common than other kinds of milk
4. cow's milk makes a better-tasting cheese than other kinds of milk

30. During the past two decades, cheese consumption in the United States has

1. decreased 2. doubled 3. tripled 4. stayed the same

31. About how much of the milk production in the United States is used to make cheese?

1. less than one-quarter
2. one-half
3. three-quarters
4. nine-tenths

Back in the early 1800's, a man named David Watkinson had a hardware store in Hartford, Connecticut. His assistants were two nephews, Samuel and David Collins. They sold iron and steel from Sheffield, England, which was used for edged tools. 149

One day, at the village blacksmith shop, David Collins heard a farmer complaining about how long it took to get a new ax. The blacksmith forged the metal into shape fast enough, he said, but he had to put on the cutting edge himself and that meant spending days at the grindstone. This gave young David an idea: "Why not make axes finished, ready to use, and sell them in our hardware store?" David Collins convinced his uncle, and within five years they were employing about forty men - all busily at work making axes.

32. The steel sold by David Watkinson came from

1. Hartford, Connecticut
2. Sheffield, England
3. the village blacksmith shop
4. the hardware store

33. David Collins worked with

1. his uncle, David Watkinson, and his brother, Samuel Collins
2. his two uncles, David Watkinson and Samuel Collins
3. his two nephews, David Watkinson and Samuel Collins
4. his nephew, David Watkinson, and his brother, Samuel Collins

34. What was the farmer's chief complaint?

1. The people who sold the steel should furnish it to the blacksmith already sharpened.
2. The metal used for the axheads was impossible to sharpen.
3. The materials available for grinding were unsatisfactory.
4. none of these

35. One of the chief advantages of David's idea was that

1. it was no longer necessary to import steel from Sheffield, England
2. it was no longer necessary for the blacksmith to forge the metal
3. people could buy an ax ready to use
4. the axes made by the new company were better than those formerly made

36. Farmers around Hartford had to grind their own axes until

1. shortly before the American Revolution
2. the early 1800's
3. five years after the first ax was made in Hartford
4. five years before steel was used for making axheads

Mount Rainier is located 50 miles southeast of the lower reaches of Puget Sound in west-central Washington. This massive mountain was once called God by the Indians who lived in the vicinity. The unmatched single-peak glacier system of Mount Rainier took centuries to build. Mount Rainier is a dormant volcano, as there is evidence that the mountain has long been asleep. However, steam jets still spout in the crater.

A dense forest clothes the slopes of Mount Rainier to about 5,000 feet. Above this point there are a few stragglers, but none above 7,000 feet, for continual snow-slides keep the surface close-shaved. Above the tree line there are meadows with flower gardens, the blooms thriving close beside the icy rivers. Many trails thread through Mount Rainier National Park; one of them, the famous Wonderland Trail, circles the mountain and requires a week on horseback to make the circuit. There are many admirable campsites, and the trout fishing is excellent.

37. In the first paragraph, the word "dormant" means

1. hot 2. large 3. holy 4. inactive

38. Today in the crater of Rainier one can see

1. Indian villages
2. steam jets
3. a few trees
4. a glacier

39. The surface of the mountain above 7,000 feet is kept bare of trees by

1. snowslides
2. lumbermen
3. trail blazers
4. icy rivers

40. Mount Rainier National Park is

1. a poor choice for campers
2. lacking in trees and flowers
3. threaded with many beautiful trails
4. of interest only because of the mountain

NAME _____ DATE OF BIRTH _____
LAST FIRST MIDDLE

SCHOOL _____ CITY _____

TEACHER _____ GRADE _____

DATE _____ FORM _____

RAW SCORE	PERCENTAGE

RAW SCORE	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
PERCENTAGE	0	3	5	8	10	13	15	18	20	23	25	28	30	33	35	38	40	43	45	48	50	53	55	58	60	63	65	68	70	73	75	78	80	83	85	88	90	93	95	98	100

MINIMUM COMPETENCE TEST IN READING

for New York State Secondary Schools

THE UNIVERSITY OF THE STATE OF NEW YORK

THE STATE EDUCATION DEPARTMENT

ALBANY

BE SURE YOUR MARKS ARE HEAVY AND BLACK.
ERASE COMPLETELY ANY ANSWER YOU WISH TO CHANGE.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4

21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4

PRINTED IN U.S.A. 10M FORM I.T.S. 1000 P 6963

APPENDIX B

READING ABILITY TEST

EXPERIMENTAL EDITION

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO.

This is a test to see how well you understand what you read. It contains 11 reading passages. You are to read each passage and answer the questions based on it. The correct answer for each question will be found among the choices following the question. Decide which choice is the correct answer. Then, on the answer sheet, blacken the answer that has the same number as the correct answer. You will have 40 minutes to answer the 40 questions.

To make sure that you know what to do, read the sample passage below and see how the questions are answered:

SAMPLE READING PASSAGE

Air has many uses. Just take a breath of the air around you. It surrounds every living thing. Kites, airplanes, birds, dust, gliders, and a number of other things are supported by the air. Even the rocket which carries men into space depends on oxygen to burn.

1. The story is mainly about

   1. breathing materials
   2. air pumps
   3. uses of air
   4. kinds of air

The correct answer is answer number 3. Therefore, a heavy black line has been drawn in the answer space under 3 on the sample answer sheet.

SAMPLE ANSWER SHEET

	1	2	3	4
1			■	
2		■		

2. Oxygen is necessary for

   1. painting
   2. burning
   3. studying
   4. dying

The correct answer is answer number 2. Therefore, a heavy black line has been drawn in the answer space under 2 on the sample answer sheet.

In the same way, on your answer sheet, answer all the questions on the reading passages. You will find the questions for each reading passage below that passage. Remember, all answers must be marked on the separate answer sheet. Blacken only one answer space for each question. Work rapidly and do the best you can.

DO NOT MAKE MARKS OF ANY KIND IN THE TEST BOOKLET.

DIRECTIONS: Read each story and answer the questions.
Mark your answers on the separate answer sheet.

153

Every kind of space capsule occasionally produces motions which are felt by the crew. Rolling, pitching, and vibrating can be felt if the astronauts push the wrong button at the wrong time. These motions can cause nausea. Tumbling in space can cause an upset stomach. This kind of sickness is called nausea. A long flight could be scrubbed if the crew continues to suffer from it. Men are very sensitive to rocking motions and can become sick easily. Nausea is not a very dangerous illness ordinarily. However, a sickness which is only uncomfortable on earth can prove hazardous in space.

1. The story says that nausea is

   1. dizziness
   2. stomach sickness
   3. rocking
   4. weakness

2. Space crews must guard against

   1. long flights
   2. fever
   3. cold weather
   4. capsule movements

3. In this story, the word "hazardous" means

   1. sensitive
   2. illness
   3. dangerous
   4. comfortable

The Daytona Speedway is not much larger than the track at Atlanta Raceway. But around that 2 1/2 mile speedway, as many as 94,000 people gather which is nearly the largest congregation of spectators in the United States. Watching under such crowded conditions usually gives most people a strong desire to get away from each other. But here, where Americans have a bird's-eye view of some of the greatest races in the world, people push to the fences unmindful of crowding each other in trying to see what is happening.

4. According to this account, races at Daytona are

   1. crowded
   2. unpleasant
   3. dangerous
   4. painful

5. The story indicates that Americans are

   1. very polite and friendly in crowds
   2. interested in what other people are doing at races
   3. anxious to get away from crowds at Daytona
   4. interested in seeing races in spite of crowded conditions

6. The story says that

   1. Daytona Speedway is somewhat larger than the Atlanta track
   2. Atlanta Speedway is larger than Daytona
   3. Atlanta is 2 1/2 miles long
   4. Both Daytona and Atlanta are the same length

Working around the used-car establishments of New York City and upper New York, Joe Johnson sold second-hand cars to anyone who desired to trade. With his price book clutched in his left hand, a pencil in his right, he would calculate figures as he bargained with everyone. Yet his concern for business did not keep him from noticing the manners, customs, clothes, and ways of life of the people he met. He found out that many customers had set out to buy vehicles with no knowledge of the problems they would encounter and that few of them had information necessary for bargaining with "car sharks." He decided it was his obligation to help other people who were planning to purchase second-hand cars and so he wrote a book entitled, Guide for Purchasers of Used Cars.

7. The story says that many of the customers whom Mr. Johnson met were

   1. poorly prepared to purchase a used-car
   2. intelligent people who bought only the best used car buys
   3. ignorant people who could not read the price book
   4. natives of New York City

8. The phrase "bargaining with" is closest in meaning to

   1. discovering
   2. trading
   3. stopping
   4. guiding

9. Mr. Johnson wrote a book because he wanted to

   1. tell people what customers are like
   2. help those who were getting ready to buy a used-car
   3. make money to expand his business
   4. spread knowledge of used-car customers among car dealers

[3]

There are places in the world where television sets are so scarce that they are almost worshipped. But whenever anything is easy to get, people generally treat it carelessly. They forget how valuable it is. In a land like ours--with both black and white, and color television--with weekly TV shows and movies, we don't think much about the television set itself. Once in a while we get upset when the set breaks down or the picture changes, wiggles, and rolls, or the color goes off. Least upset of all, however, is the owner of two sets. All he has to do is turn on the second one to get his nightly share of TV entertainment.

10. The writer is trying to

    1. remind us that a television set is more important than we seem to realize
    2. make fun of TV viewers who own two television sets
    3. tell us how foolish those people are who worship a television set
    4. warn us of the dangers connected with our excellent TV system

11. According to the writer, he lives in a land where there are

    1. black and white TV sets but no color sets
    2. frequent TV repairs
    3. both black and white sets and color sets as well
    4. only broken sets which flicker and roll

12. The writer indicates that people who experience occasional trouble with their TV sets

    1. worship TV in other lands
    2. forget that tubes blow out
    3. usually have color television
    4. are reminded that a TV set can be a problem

13. The writer is telling us that

    1. people usually appreciate their blessings
    2. modern conveniences are worshipped in our land
    3. when anything is easy to get, people tend to treat it carelessly
    4. what we get for nothing is usually worth nothing

For more years than anyone can remember, the American ideal of year-round sports has been baseball in the summer and football in the winter. Now, around the country, several other sports are becoming well-liked. One of these is indoor ice hockey. Four skills make up this fast-moving ice game. The movements of hockey are skating, stick-handling, shooting, and passing. The most important one of these moves is skating since skating is to hockey what throwing is to baseball and what running is to football. 156

Actually, hockey has been with us a very long time. It has been found only among college sports for many years. So many sports are played outdoors that indoor sports lovers flock to the hockey arena for their fun.

14. The basic movements of the game of hockey are

1. skating and passing
2. passing and shooting
3. stick-handling and shooting
4. all of these

15. Skating is to hockey what

1. running is to baseball
2. running is to football
3. throwing is to football
4. throwing is to hockey

16. Ice hockey is growing in popularity because many people

1. prefer indoor sports
2. understand hockey
3. play hockey
4. love football and baseball

It has been said that everybody talks about the Vietnam war, but nobody does much about it. Newspapers show us that the day has come when this saying is no longer true. Today we know much about war and understand its consequences better.

News media show that the fighting abroad has an amazing effect on our moods. For example, should the fighting start increasing rapidly, people at home feel restless and rebellious, while a steady slow-down in foreign battles is likely to make the population at home more relaxed, satisfied with the national condition and less likely to stage demonstrations.

As in the case of many other occurrences on the planet earth, war in all its forms, from a simple argument to a nuclear battle can often be traced to greed. Of course, other factors are involved too, but greediness is the chief cook stirring up all the ingredients for war and disunity as we know it.

17. This article may be described as

1. biographical
2. informational
3. fictional
4. humorous

18. According to the article when fighting abroad is slowed down people here at home are likely to feel

1. sad
2. nervous
3. angry
4. relaxed

19. The expression "chief cook" in the last paragraph refers to

1. war
2. the population
3. greed
4. fighting

LeRoy Keyes is probably one of the greatest athletes in football who would prefer to be recognized as the greatest player in basketball. Only O. J. Simpson can rival him as a runner. LeRoy's running ability, contrary to Simpson's is based more on maneuverability than on speed; he dodges extremely well. They are both actually giants of professional football and both competitors participated in childhood in many other challenging sports as well. Keyes and Simpson are unique players having individual styles. They manage to convey in football, as in similar activities, the portrait of the human body as a dynamic vehicle capable of surpassing equally swift competitors. 157

20. According to this story the main reason why both Keyes and Simpson are considered great football players is because of their

1. waggling
2. size
3. team spirit
4. running ability

21. The story says that Keyes would rather be

1. a basketball star
2. as short as Simpson
3. a swift competitor
4. a dynamic vehicle

22. As used in this story, the word "manage" means

1. would like 2. give orders 3. study 4. are able

The first small car for Americans was introduced in 1931 by German dealers. It was 158.7 inches long, 61 inches wide, and 59 inches tall. Forty years since it came to this country, American cars are beginning to compete with that small car. Auto companies are afraid that foreign cars in this country will outnumber our own. Today, Pinto from Ford and Vega from Chevy challenge the Volkswagen.

23. The size of the Volkswagen was

1. the same as a standard American car
2. smaller than a standard American car
3. larger than a standard American car
4. larger than all foreign cars

24. The American dealers are afraid because foreign compact cars

1. might out-number American cars
2. would very likely run well
3. are breaking speed records
4. sell for the same price as American cars

25. The story says that the first Volkswagen was

1. taller than wider
2. as high as it was wide
3. wider than taller
4. wider than longer

Pot is a mild form of narcotic. No one knows how LSD and pot differ but according to recent studies, they share accidentally the same chemical called THC. Dr. Hollister made a comparative study of the effects of the two drugs when he set out to test the reactions of each drug on several people. The same chemical in the content of the drugs given to the patients in the experiment caused effects that were both alike and different. He found that the THC caused a need for sleep but that LSD prevented sleep. 158

However, both drugs made people feel weak simply because they produced feelings which generally caused minutes to stretch into hours. All people in the experiment took the same dosage of each drug and reported their responses. People from the United States made up this experimental group. Although they did not smoke more than three to seven high-quality reefers, they felt nearly the same reaction to THC as to LSD. More than this large dose of THC may create even more similar responses. About one-tenth of each large dose was given to the patients in this experiment.

26. The central theme of this article deals with

1. how researchers make pot
2. the accidental discovery of marijuana
3. the importance of narcotics in the United States
4. comparing the effects of pot and LSD

27. When researchers investigated pot and LSD they found a similar chemical called

1. THC
2. LSD
3. marijuana
4. reefers

28. According to the article, both drugs caused

1. blindness
2. weakness
3. dizziness
4. neatness

29. According to the article, increasing the dosage of THC may

1. cause LSD users to sleep more easily
2. cause THC users to have wild dreams
3. cause effects even more like LSD
4. have no effect whatsoever

30. In comparing LSD and THC one may say that in relation to sleep, they produce effects which are

1. similar
2. opposite
3. dangerous
4. identical

31. About how much of a large dose of each drug was given to the people in this experiment?

1. approximately one-tenth
2. more than one-half
3. three-fourths
4. less than one-eighth

In the late 1900's, a man named Bert Bacharach formed and orchestra at A & R Studios, New York City. His sound engineer was Phil Ramone. They cut and sold records in the country which set the pace for new sounds in music.

One day at morning auditions, Bacharach heard a girl singing a song with a touch of soul, rock and religion in her voice. She sang the song fast enough, he thought, but she has to put more feeling into it and that meant changing the rhythm in the score. This gave Bacharach an idea: Why not write songs for the orchestra and this great singer at the same time? Bert Bacharach signed Dionne Warwick and within five years they have sold over forty hits and are still hard at work making new music.

32. The change Bacharach made in Dionne Warwick's music was in the

1. background
2. rhythm
3. tone
4. sound

33. Dionne Warwick's voice has which combination of the following qualities?

1. rock, religion, soul
2. soul, rock
3. religion, soul
4. feeling, sound, soul

34. Bacharach noticed that Dionne's voice needed

1. higher pitch
2. less expression
3. lower tone
4. more feeling

35. One of the chief advantages of Bacharach's idea was that

1. the orchestra played better with Dionne singing
2. Dionne Warwick sang to larger audiences
3. together Bert and Dionne produced many hits
4. Phil Ramone became well known

36. Bacharach heard Dionne sing for the first time

1. in his home
2. at a studio audition
3. in a theatre
4. at a nightclub

John Fitzgerald Kennedy's grave is located in the central section of Arlington National Cemetery in the State of Virginia. This circular monument is visited by thousands of people weekly. The spacious, grassy area is surrounded by a low grey-brick wall on which are engraved statements from his Inaugural Address. Here, JFK lies dormant, as quiet as the stones. However, an undying spirit still radiates from his memory.

Kennedy presided merely 1,000 days as president out of the 3,000 days he could have served in that capacity. This represents only one-third of his full term but in that brief time, he inspired people throughout the world. Much can be learned from looking into John Kennedy's life story and the same may now be said of his brother, Robert. Perhaps the greatest lesson both of them gave to the world is that none of us can afford to be only onlookers and bystanders. We all need to take action as they did.

37. In the first paragraph, the word "dormant" means

1. alone
2. happy
3. holy
4. lifeless

38. Today at Kennedy's grave one can see

1. John Kennedy's picture
2. JFK's words carved in stone
3. an American flag
4. all of Arlington Cemetery

39. The large monument over Kennedy's grave is in the shape of a

1. circle
2. square
3. triangle
4. rectangle

40. Arlington National Cemetery is

1. a poor choice for Kennedy's grave
2. lacking in trees and flowers
3. crowded weekly with many visitors
4. of interest only to those who have relatives buried there

NAME _____ DATE OF BIRTH _____
LAST FIRST MIDDLE

SCHOOL _____ CITY _____

TEACHER _____ GRADE _____

DATE _____ FORM _____

RAW SCORE | PERCENTAGE

RAW SCORE	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
PERCENTAGE	0	3	5	8	10	13	15	18	20	23	25	28	30	33	35	38	40	43	45	48	50	53	55	58	60	63	65	68	70	73	75	78	80	83	85	88	90	93	95	98	100

READING ABILITY TEST

CITY BOARD OF EDUCATION

ALBANY, NEW YORK

BE SURE YOUR MARKS ARE HEAVY AND BLACK.
ERASE COMPLETELY ANY ANSWER YOU WISH TO CHANGE.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40		

APPENDIX C

ANSWER KEY TO THE NEW YORK STATE MINIMUM COMPETENCE TEST AND THE READING ABILITY TEST

Paragraph	Item Number	Number of Correct Answers
A	1	2
	2	4
	3	3
B	4	1
	5	4
	6	1
C	7	1
	8	2
	9	2
D	10	1
	11	3
	12	4
	13	3
E	14	4
	15	2
	16	1
F	17	2
	18	4
	19	3
G	20	4
	21	1
	22	4
H	23	2
	24	1
	25	3

Paragraph	Item Number	Number of Correct Answers
I	26	4
	27	1
	28	2
	29	3
	30	2
	31	1
J	32	2
	33	1
	34	4
	35	3
	36	2
K	37	4
	38	2
	39	1
	40	3

APPENDIX D

DIRECTIONS FOR ADMINISTRATION OF THE NEW YORK STATE MINIMUM COMPETENCE TEST IN READING AND THE READING ABILITY TEST

The following directions are to be used for administering the tests with IBM 805 answer sheets. The underlined material should be read verbatim.

After the desks have been cleared of books and papers and the pupils are ready to begin the test, read the following:

Today you are going to take two tests which measure how well you read. I will pass out the first test booklets with the answer sheets. Leave the test booklet on your desk. Do not open the booklet until you are told to do so. While you are waiting, you may fill in the heading of your answer sheet.

Distribute the test booklets and answer sheets, face up, one to each pupil. Be sure that one half of the group to be tested receives Minimum Competence Test booklets and answer sheets; the second half of the group receives Reading Ability Test booklets and answer sheets. After the rest period, tests will be interchanged.

When pupils have completed filling in the heading of the answer sheets, say:

Now read the directions on the front of your test booklet, but do not open the booklet until I tell you to do so.

When pupils have finished, say:

Does everyone know how to mark the answer sheet? Are there any questions? (Pause)

If the pupils do not understand the sample questions, the appropriate directions and explanations may be read again or paraphrased. However, the teacher should not undertake to provide an independent explanation or a new sample. When directions are clear, say:

Now open the booklet and begin work.

Make a written note of the time the pupils begin the test. Walk around the room to make sure that the pupils understand the directions and are marking answers properly.

After exactly 40 minutes of working time say:

Stop working. Put your pencils down. Close your booklets so that the cover page is on top. (Pause)

Collect answer sheets and booklets, say:

You may stand up and stretch for five minutes.

At the end of five minutes, say:

Please be seated and you will take the second test which is answered in the same way. I will pass out the booklets and answer sheets. Please do not open the booklet until you are given the signal but you may fill in the answer sheet heading as before.

Pass out booklets and answer sheets reversing the order of testing with each half of the group. When headings are completed, say:

As you did before, read the directions on the front of your booklet but do not start until I tell you. Are there any questions? (Pause)

You may begin.

As before, note the time pupils begin the test. Check to see if answer sheets are being filled in properly and at the end of exactly 40 minutes say:

Stop working. Put your pencils down. Close booklets so that the cover page is on top.

Collect the test booklets and answer sheets. Keep answer sheets separated and mark them to identify which groups took which test first.

APPENDIX E

LISTING OF UNFAMILIAR WORDS IN THE MINIMUM COMPETENCE TEST IN READING AND THE READING ABILITY TEST DERIVED FROM THE DALE-CHALL LIST OF FAMILIAR WORDS

Paragraph		Minimum Competence Test	Reading Ability Test
A.	1.	occasionally	capsule
	2.	produces	occasionally
	3.	offspring	produces
	4.	lobsters	motions
	5.	giraffes	crew
	6.	creatures	vibrating
	7.	albinos	astronauts
	8.	human	motions
	9.	albinos	nausea
	10.	tribes	tumbling
	11.	albinos	stomach
	12.	human	nausea
	13.	albino	continues
	14.	sensitive	sensitive
	15.	easily	motions
	16.	albino	easily
	17.	actually	nausea
	18.	merely	illness
	19.	lacks	ordinarily
	20.	pigment	hazardous
	21.	coloring	hazardous
B.	1.	total	entire
	2.	population	colorful
	3.	conditions	endless
	4.	naturally	ragged
	5.	prevent	beyond
C.	1.	crude	establishment
	2.	preached	desired
	3.	Gospel	clutched
	4.	propped	calculate
	5.	Sermons	concern
	6.	concern	manners
	7.	religion	customs
	8.	noticing	vehicles
	9.	manners	knowledge
	10.	customs	problems

Paragraph		Minimum Competence Test	Reading Ability Test
	11.	pioneers	encounter
	12.	knowledge	information
	13.	problems	necessary
	14.	equipment	bargaining
	15.	capable	sharks
	16.	coping	obligation
	17.	wilderness	purchase
	18.	pioneers	entitled
	19.	guide	guide
	20.	emigrants	purchasers
D.	1.	scarce	television
	2.	worshipped	scarce
	3.	usually	worshipped
	4.	concerned	valuable
	5.	concerned	television
	6.	dweller	television
	7.	usually	entertainment
E.	1.	required	hockey
	2.	actually	movements
	3.	revolution	hockey
	4.	reputation	skills
	5.	enviable	hockey
	6.	slums	actually
	7.	fled	hockey
	8.	suburbs	hockey
F.	1.	scientists	consequences
	2.	studies	media
	3.	amazing	abroad
	4.	effect	amazing
	5.	moods	effect
	6.	dispositions	moods
	7.	creativity	example
	8.	example	increasing
	9.	barometer	restless
	10.	steadily	rebellious
	11.	restless	steady
	12.	irritable	foreign
	13.	pessimistic	population
	14.	steady	relaxed
	15.	barometric	satisfied
	16.	pressure	national
	17.	apt	condition
	18.	relaxed	demonstrations
	19.	satisfied	occurrence
	20.	occurrences	argument
	21.	planet	nuclear
	22.	tornado	greed

Paragraph		Minimum Competence Test	Reading Ability Test
	23.	Old Sol	factors
	24.	atmosphere	involved
	25.	involved	greed
	26.	chef	stirring
	27.	ingredients	ingredients
	28.	climate	disunity
G.	1.	probably	probably
	2.	architecture	athletes
	3.	consider	prefer
	4.	primarily	recognized
	5.	architect	rival
	6.	rival	ability
	7.	sculptor	contrary
	8.	builder	maneuverability
	9.	architectural	dodges
	10.	tradition	extremely
	11.	gigantic	actually
	12.	sculpture	professional
	13.	previous	competitors
	14.	traditions	participated
	15.	historical	challenging
	16.	architecture	unique
	17.	transformed	individual
	18.	unique	styles
	19.	creation	manage
	20.	revision	convey
	21.	improvement	similar
	22.	traditional	activities
	23.	styles	portrait
	24.	managed	human
	25.	convey	dynamic
	26.	architecture	vehicle
	27.	sculpture	capable
	28.	human	surpassing
	29.	confines	competitors
H.	1.	colonies	introduced
	2.	press	dealers
	3.	illegal	compete
	4.	governor	companies
	5.	press	foreign
	6.	democracy	challenge
I.	1.	product	narcotic
	2.	according	according
	3.	ancient	recant
	4.	legend	accidentally
	5.	accidentally	chemical
	6.	merchant	comparative

Paragraph		Minimum Competence Test	Reading Ability Test
	7.	supply	effects
	8.	pouch	reactions
	9.	stomach	chemical
	10.	rennet	content
	11.	lining	patients
	12.	pouch	experiment
	13.	combined	effects
	14.	curd	prevented
	15.	whey	produced
	16.	whey	experiment
	17.	satisfied	experiment
	18.	satisfied	responses
	19.	simply	high-quality
	20.	exception	reefers
	21.	although	reaction
	22.	billion	dose
	23.	one-tenth	similar
	24.	annual	responses
	25.	annual	patients
	26.	production	experiment
J.	1.	assistants	orchestra
	2.	nephews	studios
	3.	complaining	engineer
	4.	forged	auditions
	5.	convinced	religion
	6.	employing	rhythm
	7.	busily	orchestra
K.	1.	located	located
	2.	massive	central
	3.	vicinity	section
	4.	unmatched	National
	5.	glacier	circular
	6.	system	monument
	7.	centuries	spacious
	8.	dormant	grassy
	9.	volcano	area
	10.	evidence	surrounded
	11.	jets	engraved
	12.	spout	statements
	13.	crater	Inaugural
	14.	dense	dormant
	15.	slopes	undying
	16.	stragglers	spirit
	17.	continual	radiates
	18.	thriving	memory
	19.	trails	presided
	20.	famous	president

Paragraph	Minimum Competence Test	Reading Ability Test
21.	trail	capacity
22.	requires	represents
23.	circuit	inspired
24.	admirable	afford
25.	campsites	onlookers
26.	trout	bystanders
27.	excellent	action

APPENDIX F

RAW SCORES OF 151 SUBJECTS ON THE NEW YORK STATE MINIMUM COMPETENCE TEST IN READING AND THE READING ABILITY TEST

Minimum Competence Test	Reading Ability Test	Minimum Competence Test	Reading Ability Test
18	9	27	21
24	9	14	22
14	10	17	22
17	12	20	22
22	12	12	23
22	13	20	23
14	14	16	23
32	15	18	24
18	16	19	24
15	16	31	25
23	16	21	25
18	16	24	25
15	17	29	25
20	17	29	25
17	17	20	26
7	18	23	26
19	18	25	26
12	19	24	26
14	19	20	27
15	19	24	27
20	20	32	27
12	20	31	27
27	20	20	27
27	21	13	27
20	21	29	28
17	21	29	28

Minimum Competence Test	Reading Ability Test	Minimum Competence Test	Reading Ability Test
25	28	30	33
27	28	21	33
15	28	34	33
29	28	33	33
23	29	34	33
33	29	32	34
28	29	32	34
31	29	34	34
29	29	24	34
26	30	27	34
31	30	36	34
27	30	34	34
18	30	32	35
25	31	35	35
28	31	31	35
31	31	30	35
34	31	33	35
25	31	35	35
33	31	32	35
33	31	26	35
25	32	31	36
25	32	37	36
28	32	33	36
28	32	37	36
24	32	35	36
20	32	35	36
32	32	33	36
27	32	35	36
28	33	38	36
32	33	38	36
33	33	32	36
32	33	35	36
38	33	36	37

Minimum Competence Test	Reading Ability Test
32	37
36	37
30	37
35	37
28	37
33	37
37	37
35	37
37	38
32	38
30	38
35	38
35	38
32	38
39	38
38	38
32	38
36	38
38	38
36	38
29	38
34	38
37	39
32	39
36	39
38	39
34	39
39	39
37	39
39	39
39	40
38	40
38	40

Minimum Competence Test		Reading Ability Test	
N	= 151	N	= 151
Σy_2	= 4225	Σx_2	= 4504
Σy^2	= 126833	Σx^2	= 143454
M	= 27.7960	M	= 29.8278
SD	= 7.8877	SD	= 7.7929
SE_M	= .6397	SE_M	= .6341

APPENDIX G

ITEM ANALYSIS DATA FOR HIGH (H) AND LOW (L) SCORERS

MINIMUM COMPETENCE TEST

Item No.	H-L	Number Choosing A	B	C	D	O	Percentage Right	
1	H	1	67	8			88	78
	L	4	50	12	9		67	
2	H		1		75		99	94
	L	4	1	4	66		88	
3	H			75	1		99	92
	L	1	3	63	8		84	
4	H	74		1	1		97	92
	L	65	5	3	2		87	
5	H		4	2	70		92	74
	L	5	24	4	42		56	
6	H	64		12	1		83	72
	L	46	7	13	9		61	
7	H	52	4	12	8		68	58
	L	36	19	8	12		48	
8	H	2	70	3	1		92	70
	L	16	36	9	14		48	
9	H	3	72	1			95	79
	L	15	47	4	9		63	
10	H	70	2		4		92	84
	L	63	3	6	9		76	
11	H	1	14	57	4		75	60
	L	5	24	33	13		44	
12	H	3	6	6	60		80	64
	L	7	16	16	35	1	47	

Item No.	H-L	Number Choosing A	B	C	D	O	Percentage Right	
13	H	1	28	47			62	52
	L	4	30	31	10		41	
14	H		1	2	73		96	85
	L	1	5	14	55		73	
15	H	5	67	2	1		89	79
	L	13	52	3	7		69	
16	H	74	1		1		97	87
	L	57	5	5	7	1	76	
17	H	13	58	4	1		76	63
	L	25	46	11	1	1	49	
18	H	2		3	71		93	79
	L	6	8	12	49		65	
19	H	3	4	64	5		84	64
	L	8	11	33	23		44	
20	H	5	4	13	54		71	59
	L	12	15	13	38		47	
21	H	73	2				97	83
	L	55	19	5	7	1	69	
22	H	1	1	3	71		93	81
	L	8	8	7	52		69	
23	H	2	66	5	3		87	74
	L	3	45	13	14		60	
24	H	57	8	1	9		76	62
	L	35	17	6	16	1	47	
25	H	12	1	68			83	78
	L	5	6	49	8	2	72	
26	H	4	38	6	28		37	27
	L	14	29	17	12	3	16	
27	H	62	3	11			82	67
	L	38	9	19	7	2	51	
28	H	12	60	1	3		79	54
	L	29	22	8	14	2	29	
29	H		13	62	1		82	57
	L	8	29	24	12	2	32	

Item No.	H-L	Number Choosing A	B	C	D	O	Percentage Right	
30	H		75		1		99	74
	L	12	37	19	5	2	49	
31	H	68	4	2	2		89	69
	L	37	13	8	15	2	49	
32	H	8	67	1			88	69
	L	26	37	5	4	3	49	
33	H	69	2	5			91	76
	L	45	7	12	7	4	60	
34	H	6	1	1	68		89	71
	L	11	13	8	40	3	53	
35	H	3	2	70	1		92	74
	L	4	14	41	9	7	55	
36	H	1	50	16	8		67	47
	L	9	21	23	16	7	27	
37	H		6	4	66		87	60
	L	17	24	1	25	8	33	
38	H	2	61	7	6		80	63
	L	13	29	14	5	9	45	
39	H	72	2		2		95	77
	L	33	5	5	12	9	59	
40	H			70	6		92	73
	L	1	5	39	19	10	53	

READING ABILITY TEST

1	H		73	2	1		92	83
	L	1	63	8	3		84	
2	H	9		3	64		84	72
	L	25	4	2	44		59	
3	H	1	2	71	2		93	77
	L	2	20	46	6	1	61	
4	H	75	1				99	92
	L	63	3	8	1		84	

Item No.	H-L	Number Choosing A	B	C	D	O	Percentage Right	
5	H			6	70		92	73
	L	6	14	15	40		53	
6	H	59	6	6	6		76	61
	L	33	14	16	11		45	
7	H	70	1	2	2		93	75
	L	35	26	2	5		56	
8	H	1	69	4	2		91	84
	L	8	57	5	5		76	
9	H		76				100	82
	L	13	47	2	13		63	
10	H	72	1	3			95	78
	L	46	7	15	7		61	
11	H	1	5	67	3		88	78
	L	10	9	51	5		68	
12	H	1	8		67		88	72
	L	11	11	12	41		55	
13	H	1	5	70			92	78
	L	10	11	48	6		64	
14	H	5		2	67		91	82
	L	10	5	5	57		73	
15	H	4	70	2			92	76
	L	11	45	11	7	1	60	
16	H	69	3	4			91	78
	L	48	11	5	11		64	
17	H	7	65	1	1		88	70
	L	15	39	13	7	2	51	
18	H	1	6	8	61		80	62
	L	14	15	13	32	1	43	
19	H	3	13	43	17		57	42
	L	21	5	20	27	2	27	
20	H		3	3	70		92	77
	L	6	6	17	46		61	
21	H	71		2	3		93	73
	L	40	7	15	13		53	

Item No.	H-L	Number Choosing A	B	C	D	O	Percentage Right	
22	H		1	1	74		97	81
	L	12	8	6	49		65	
23	H		75	1			99	79
	L	6	44	8	17		59	
24	H	74			2		97	88
	L	59	4	6	6		79	
25	H	11	7	54	3		72	64
	L	14	17	42	3		55	
26	H	1	4	5	65		87	76
	L	3	11	11	49	3	65	
27	H	75	1				97	87
	L	57	7	6	3	2	76	
28	H	1	74	1			97	81
	L	10	49	12	2	2	65	
29	H	6	10	60			79	53
	L	19	15	28	11	2	37	
30	H	16	55	5			72	50
	L	26	21	17	9	2	28	
31	H	69	2	5			91	76
	L	46	14	6	7	2	61	
32	H	1	72	2	3		92	77
	L	2	44	9	15	3	61	
33	H	73	1	1	1		96	82
	L	50	9	6	8	2	67	
34	H	3	1		73		95	84
	L	4	7	6	54	3	73	
35	H	5	1	69			92	73
	L	20	8	40	5	2	53	
36	H		72	4	1		93	85
	L	6	49	8	8	3	77	
37	H	10	1	6	60		78	55
	L	16	16	15	23	5	31	

Item No.	H-L	Number Chooseing A	B	C	D	O	Percentage Right	
38	H		72	3	1		95	79
	L	3	47	11	9	5	63	
39	H	64	3	3	6		84	61
	L	28	20	12	8	7	37	
40	H		2	72	2		95	78
	L	4	5	45	17	4	60	

APPENDIX H

SAMPLE VALUE TOOLS UTILIZED BY TEACHER IN THE DETERMINATION OF ACTUAL PUPIL INTERESTS

Thirty Clarifying Responses[1]

1. Is this something that you prize?
2. Are you glad about that?
3. How did you feel when that happened?
4. Did you consider any alternative?
5. Have you felt this way for a long time?
6. Was that something that you yourself selected or chose?
7. Did you <u>have</u> to choose that; was it a free choice?
8. Do you <u>do</u> anything about that idea?
9. Can you give me some examples of that idea?
10. What do you mean by _____: can you define that word?
11. Where would that idea lead; what would be its consequences?
12. Would you really <u>do</u> that or are you just talking?
13. Are you saying that...(repeat the statement)?
14. Did you say that...(repeat in some distorted way)?
15. Have you thought much about that idea (or behavior)?
16. What are some good things about that notion?
17. What do we have to assume for things to work out that way?
18. Is what you express consistent with...(Note something else the person said or did that may point to an inconsistency)?
19. What other possibilities are there?
20. Is that a personal preference or do you think most people should believe in that?
21. How can I help you do something about your idea?
22. Is there a purpose back of this activity?
23. Is that very important to you?
24. Do you do this often?
25. Would you like to tell others about your idea?
26. Do you have any reasons for saying or doing that?
27. Would you do the same thing over again?
28. How do you know it's right?
29. Do you value that?
30. Do you think people will always believe that?

---

[1]Louis E. Raths and others, <u>Values and Teaching</u> (<u>C</u>) 1966, p. 230. Reprinted by permission of Louis E. Raths, State University College at Fredonia, New York.

Twenty-one Clarifying Strategies[1]

1. The clarifying response
2. The value sheet
3. The value-clarifying discussion
4. Role-playing
5. The contrived incident
6. Zig-Zag lessons
7. Devil's advocate
8. Value continuum
9. Thought sheets
10. Weekly reaction sheets
11. Open-ended questions
12. Coded student papers
13. Time diaries
14. Autobiographical questionnaires
15. Public interviews
16. Decision-making interviews
17. Voting
18. Five-minute quotes without comment
19. Student reports
20. Action projects
21. An approach to self-conception

---

[1]Raths and others, p. 231.

APPENDIX I

RATING SHEET

_____
Name

Please rate, by the use of check marks, each of the paragraphs on the Reading Ability Test as they compare with each of the paragraphs on the New York State Minimum Competence Test in Reading. Paragraphs were designed to be "congruent with" those on the New York Test; congruent refers to the similarity of format, grammatical structure, flow of thought, readability level,* and pattern of questioning. Note that relevant content was substituted within a similar framework in the place of less relevant content; please indicate the degree to which you think this was accomplished for each paragraph on the Reading Ability Test.

Paragraph	Very Good	Acceptable	Acceptable with Modification (specify on back)	Question-able	Unaccept-able
A					
B					
C					
D					
E					
F					
G					
H					
I					
J					
K					

*The Dale-Chall Formula of Readability[1] has been applied to each paragraph on both tests and statistically, matching paragraphs are equal in difficulty.

[1]Edgar Dale and Jeanne S. Chall, "A Formula for Predicting Readability," Educational Research Bulletin, I (1962), 1-9, 11-28.